

BOGDAN VOICU, OVIDIU VOICU

EVALUAREA IMPACTULUI DATA.GOV.RO ȘI A DESCHIDERII SPRE OPEN DATA ÎN ROMÂNIA



PRESA UNIVERSITARĂ CLUJEANĂ



EVALUAREA IMPACTULUI data.gov.ro ȘI A DESCHIDERII SPRE OPEN DATA ÎN ROMÂNIA

Bogdan Voicu, Ovidiu Voicu



Referenți științifici:

Prof. univ. dr. Mircea Comșa

Conf. univ. dr. Marian Vasile

Fotografie copertă: (c) Bogdan Voicu

ISBN 978-606-37-1852-6

© 2023 Autorii volumului. Toate drepturile rezervate. Reproducerea integrală sau parțială a textului, prin orice mijloace, fără acordul autorilor, este interzisă și se pedepsește conform legii.

Universitatea Babeș-Bolyai
Presa Universitară Clujeană
Director: Codruța Săcelean
Str. Hasdeu nr. 51
400371 Cluj-Napoca, România
Tel./fax: (+40)-264-597.401
E-mail: editura@editura.ubbcluj.ro
<http://www.editura.ubbcluj.ro/>

Volumul de față este realizat de Bogdan Voicu și Ovidiu Voicu în urma unui contract realizat de Romanian Association for Quantitative Studies (RQSA) pentru Guvernul României. Volumul aduce la un loc texte scrise pentru rapoartele produse în varii etape ale proiectului menționat. Opiniile exprimate în volum aparțin autorilor și nu implică în vreun fel organizațiile ce au fost implicate în proiect.

Autorii mulțumesc Larisei Barac pentru citirea atentă și comentarii pe text.

Despre autori:

Bogdan Voicu este sociolog, cercetător 1 la Institutul de Cercetare a Calității Vieții al Academiei Române și profesor de sociologie la Universitatea „Lucian Blaga” din Sibiu. Este membru fondator al RQSA. Detalii asupra publicațiilor și activității sale academice și de consultanță pot fi regăsite pe site-ul propriu: <http://web.BogdanVoicu.ro>.

Ovidiu Voicu este politolog, director executiv și fondator al Centrului pentru Inovare Publică, membru fondator al RQSA. În cadrul Centrului pentru Inovare Publică, coordonează programele privind guvernarea deschisă și drepturile fundamentale.

RQSA (Romanian Quantitative Studies Association) este o asociație al cărei scop principal este promovarea utilizării de date cantitative în științele sociale din România, incluzând utilizarea de metode cantitative la standarde contemporane, analiza cantitativă a big data (inclusiv big data constituit din corpus de natură calitativă) și deschiderea datelor etc. O parte din proiecte sunt disponibile pe www.rqsa.ro, acolo unde RQSA distribuie și seturi de date în regim deschis.

CUPRINS PE SCURT: PRINCIPALELE CAPITOLE

Introducere	12
Optiuni metodologice	14
EXEMPLE DE ANALIZE DIN ALTE ȚĂRI	14
PRINCIPIILE EVALUĂRII IMPACTULUI, LA MODUL GENERAL.....	16
PRINCIPIILE METODOLOGIEI PROPUSE	17
METODE DE COLECTARE A DATELOR	19
Rezultate	22
ANALIZA ACCESĂRIILOR SITE-ULUI ȘI A STRUCTURII DATA.GOV.RO	22
STUDIILE DE CAZ.....	28
PANELUL DELPHI	43
SONDAJELE	52
SATISFAȚIA CU DATA.GOV.RO	87
ANALIZA COMPARATIVĂ A IMPACTULUI	91
Concluzii asupra impactului.....	93
Recomandări.....	97
DIRECȚII DE ACȚIUNE	97
RECOMANDĂRI DE PROGRAME	97
INSTRUMENTE DE MONITORIZARE	100
Referinte.....	102
Anexe	105
INSTRUMENTE PROPUSE PENTRU MONITORIZAREA DATA.GOV.RO	105
CE SETURI NOI DE DATE SUNT SOLICITATE PE DATA.GOV.RO	111

Introducere	12
Opțiuni metodologice	14
EXEMPLE DE ANALIZE DIN ALTE ȚĂRI	14
PRINCIPIILE EVALUĂRII IMPACTULUI, LA MODUL GENERAL.....	16
PRINCIPIILE METODOLOGIEI PROPUSE	17
METODE DE COLECTARE A DATELOR	19
Rezultate	22
ANALIZA ACCESĂRILOR SITE-ULUI ȘI A STRUCTURII DATA.GOV.RO	22
Back-log analysis.....	22
Structura site-ului data.gov.ro.....	23
STUDIILE DE CAZ.....	28
Reutilizatori ai data.gov.ro	29
R1: servicii de location intelligence	29
I. Ce face R1	29
II. Date utilizate și produse	29
III. Relația cu data.gov.ro	29
IV. Valoarea adăugată a proiectului.....	30
R2 și utilizarea banilor publici.....	30
I. Ce face R2	30
II. Date utilizate și produse	30
III. Relația cu data.gov.ro	31
IV. Valoarea adăugată a proiectului.....	31
R3 – Bonitatea companiilor	32
I. Ce face R3	32
II. Date utilizate și produse	32
III. Relația cu data.gov.ro	32
IV. Valoarea adăugată a proiectului.....	33
R4 – despre meserii și orientare școlară	33
I. Ce face R4	33
II. Date utilizate și produse	33
III. Relația cu data.gov.ro	34
IV. Valoarea adăugată a proiectului.....	35

Non-utilizatori: U1, Auditare financiară	35
<i>I. Ce face U1</i>	35
<i>II. Date utilizate și produse</i>	35
<i>III. Relația cu data.gov.ro</i>	36
<i>IV. Valoarea adăugată a proiectului</i>	36
Utilizatori academici: concretizarea în publicații	36
Concluzii.....	40
Tipuri de impact.....	40
Puncte tari: disponibilitate, notorietate și legitimitate	40
Puncte slabe: actualizare, completitudine, extindere	41
Limitări: calitate și prelucrare automată, API, feed-back.....	41
Așteptări: noi seturi de date.....	42
PANELUL DELPHI.....	43
Metoda	43
Rate de răspuns	45
Tipuri de reacții.....	45
Rezultatele.....	46
Notorietatea	46
Legitimitatea.....	48
Disponibilitatea (oferta de date deschise)	49
Utilizarea și Re-Utilizarea.....	50
Concluzii și implicații.....	51
SONDAJELE	52
Metoda	52
Notorietatea	53
Ratele de răspuns	53
Cunoașterea.....	55
<i>Tipuri de date deschise</i>	55
<i>Caracteristici ale datelor deschise</i>	57
<i>Surse ale cunoașterii</i>	60
<i>România vs restul lumii: reprezentări subiective</i>	61
<i>data.gov.ro</i>	61
Legitimitatea.....	63
Moduri de deschidere a datelor: mediul academic	63
Mediul academic: Condiții de deschidere a datelor	64

Acceptarea reutilizării.....	66
Opinii despre utilitatea datelor deschise.....	69
Disponibilitatea.....	70
Constrângeri și catalizatori ai deschiderii datelor	70
Costuri, gratuitate și restricționări ale accesului.....	72
Responsabilități și proceduri în deschiderea și gestiunea datelor	74
Nevoi de formare.....	76
Recuperarea seturilor de date vechi și datele produse de terți.....	77
Open Data Directive	79
Utilizare și impact	80
Incidența utilizării de date printre respondenți	80
<i>Administrația Publică</i>	80
<i>Mediul academic: producție de date</i>	82
<i>Mediul academic: Utilizare și reutilizare</i>	82
Reprezentări despre consecințele deschiderii datelor.....	83
<i>Deschiderea datelor ca proces evolutiv?</i>	83
<i>Mediul academic: consecințe ale datelor deschise asupra integrității academice</i>	84
data.gov.ro	85
SATISFAȚIA CU DATA.GOV.RO	87
Rezultate din sondaje	87
Rezultate din chestionarul de satisfacție implementat pe site	88
ANALIZA COMPARATIVĂ A IMPACTULUI	91
<u>Concluzii asupra impactului.....</u>	<u>93</u>
<u>Recomandări.....</u>	<u>97</u>
DIRECȚII DE ACȚIUNE	97
RECOMANDĂRI DE PROGRAME	97
Dezvoltarea platformei și creșterea calității datelor	97
data.gov.ro ca element integrator	98
Difuzarea de date produse în mediul academic	98
Program de popularizare	99
Program de monitorizare permanentă	99
Data.gov.ro ca furnizor de pregătire și cunoaștere.....	99
INSTRUMENTE DE MONITORIZARE	100
Chestionare de feedback asociat data.gov.ro	100

Back-log și almetrics	101
<u>Referinte</u>	<u>102</u>
<u>Anexe</u>	<u>105</u>
INSTRUMENTE PROPUSE PENTRU MONITORIZAREA DATA.GOV.RO	105
Chestionar de satisfacție a utilizării data.gov.ro	105
Chestionar de feedback după 3 săptămâni	109
CE SETURI NOI DE DATE SUNT SOLICITATE PE DATA.GOV.RO	111

Figura 1. Utilizatori pe lună noiembrie 2016 – februarie 2021	22
Figura 2. Durata medie a sesiunii: februarie 2015 – ianuarie 2021, medii pe zile	22
Figura 3. Pagini pe sesiune, medii săptămânale, februarie 2018-februarie 2020.....	23
Figura 4. Capturi de ecran a paginii de intrare din câteva portaluri	25
Figura 5. Topologia discuției Delphi: cu albastru sunt temele inițiale, cu verde sunt argumente pro, iar cu roșu argumente contra.....	45
Figura 6. Structura arborescentă a răspunsurilor, cu detalieri pe ramura „notorietății”	47
Figura 7. Opinii despre datele produse în organizația respondentului	55
Figura 8. Medii ale indicatorului de cunoaștere (cu estimări ale intervalelor de încredere 95%)	59
Figura 9. Surse de informare despre open data	60
Figura 10. Notorietatea data.gov.ro în cele trei eșantioane	62
Figura 11. Localizarea data.gov.ro.....	63
Figura 12. Mediul academic: În ce condiții vă simțiți confortabil să lăsați libere datele produse de dvs.	65
Figura 13. Mediul academic: Ce probleme/rețineri aveți legate de a deschide datele produse de dvs.	66
Figura 14. Cunoaștere despre date și deschidere către reutilizare	68
Figura 15. Opinii despre beneficiile deschiderii datelor	70
Figura 16. Opinii despre factorii ce pot facilita deschiderea datelor în organizația de apartenență.....	72
Figura 17. Responsabilități în deschiderea datelor	74
Figura 18. Proceduri de gestiune a datelor în administrația publică	75
Figura 19. Preocupări privind deschiderea datelor mai vechi	77
Figura 20. Opinii despre folosirea terților în procesul de deschidere a datelor.....	78
Figura 21, Opinii ale administrației publice legate de implementarea Open Data Directive.....	79
Figura 22. Probabilitatea de a utiliza date de pe data.gov.ro în funcție de nivelul de cunoaștere a ce înseamnă date deschise	81
Figura 23. Tipuri de date folosite în mediul academic	82
Figura 24. Utilizarea data.gov.ro de către respondenții I sondajele din 2020	85
Figura 25. Publicarea de date pe data.gov.ro.....	86
Figura 26. Satisfacția cu diferite elemente ale data.gov.ro.....	87
Figura 27. Număr de completări pe zile ale chestionarului de satisfacție disponibil pe data.gov.ro în perioada 20 decembrie 2020-4 martie 2021.....	88
Figura 28. Detalii privind cei ce au completat chestionarul de satisfacție	89
Figura 29. Satisfacția utilizatorilor data.gov.ro în funcție de sectorul de activitate	90
Figura 30. Evoluția Opend Mariturity Data între 2016 și 2020 în câteva țări europene	91
Figura 31. Impactul creșterii scorului Open Data Maturity asupra creșterii PIB/locuitor.....	92

Tabelul 1. Principiile metodologiei	17
Tabelul 2. Sinteza metodelor folosite pentru colectarea datelor	19
Tabelul 3. Statistici pentru perioada februarie 2018 – februarie 2021.....	23
Tabelul 4. Trecere în revistă a câtorva caracteristici ale site-urilor portalurilor guvernamentale de date deschise din câteva țări europene	24
Tabelul 5. Lista studiilor de caz.....	28
Tabelul 6. Impactul academic și în literatura gri a portalurilor de date deschise din EU, țările asociate și Moldova.....	37
Tabelul 7. Tabel sinoptic al analizei asupra articolelor luate în considerare.....	39
Tabelul 8. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Notorietatea	46
Tabelul 9. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Notorietatea	48
Tabelul 10. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Notorietatea	49
Tabelul 11. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Utilizarea și re-utilizarea	50
Tabelul 12. Număr de respondenți în funcție de poziționarea față de sondaje	54
Tabelul 13. Categoriile de personal care au tendința de a indica mai des decât ceilalți anumite caracteristici ce fac datele să fie date	56
Tabelul 14. Reprezentări despre caracteristicile datelor deschise („În ce măsură fiecare dintre următoarele elemente credeți că se aplică datelor deschise (open data)”).....	57
Tabelul 15. Structura eșantioanelor conform indicatorilor de cunoaștere calculați	59
Tabelul 16. Evaluarea subiectivă a publicării de date deschise în România și în UE	61
Tabelul 17. Modalități considerate legitime de a deschide datele: mediul academic, 2020	64
Tabelul 18. Reprezentări ale administrației publice în ce privește legitimitatea (re)utilizării datelor deschise în funcție de scopul (re)utilizării	67
Tabelul 19. Deschiderea spre utilizarea datelor deschise	67
Tabelul 20. Costul utilizării datelor deschise	69
Tabelul 21. Opinii despre factorii ce pot contribui la deschiderea datelor	71
Tabelul 22. Opinii despre eventuala monetizare sau deschidere restricționată a datelor publice	73
Tabelul 23. Opinii despre guvernarea datelor deschise	76
Tabelul 24. Dvs. credeți că ați avea nevoie de cursuri certificate în deschiderea datelor?	76
Tabelul 25. Utilizarea de date la nivelul eșantionului de angajați din administrația publică din 2020.....	80
Tabelul 26. Opinii privind consecințele deschiderii datelor: 2020	83
Tabelul 27 Satisfacția cu data.gov.ro.....	90
Tabelul 28. Tabel sinoptic al principalelor rezultate	94

Caseta 1. Lista articolelor academice studiate	38
Caseta 2. Prezentarea discuției așa cum a fost ea afișată pe Kialo experților invitați ca participanți	44
Caseta 3. Exemplu de mesaj de refuz de a răspunde	53
Caseta 4. Modul de construcție a indicatorului de cunoaștere a datelor deschise	58

Odată cu dezvoltarea capacității de calcul dată de progresul tehnologiei IT, civilizația umană a început să redescopere și să redefinească moduri raționale de organizare a sa. Mai exact, în efortul de informare a acțiunilor umane, folosirea datelor a căpătat noi valențe. Oamenii pot folosi date din trecut, pentru a documenta efectele diferitelor acțiuni întreprinse, pentru a cunoaște mai bine situația prezentă, pentru a putea prezice ce consecințe vor avea inițiative individuale sau colective, indiferent că acest lucru se petrece în viața de zi cu zi, personală sau de business, din sectorul public sau din organizații non-profit.

Oamenii generează non-stop date, iar o bună parte din ele sunt plasate în sectorul guvernamental sau, la modul și mai general, sunt generate din bani publici. Aceste date au azi capacitatea de a fi circulate, analizate și să devină un element esențial de input în orice sferă a vieții.

Fără a intra în detalii despre filosofia și utilitatea datelor, schițate extrem de sumar în paragrafele anterioare, este util să notăm faptul că principiul a fost formalizat în spațiul public prin manifestul *Open Data Charter* publicat de G8 – grupul celor mai dezvoltate economii – în anul 2013 (G8, 2013). Manifestul a fost elaborat prin dezvoltarea unei platforme cooperative reunind azi peste 100 de guverne și organizații transnaționale¹ și a fost revizuit căpătând o nouă concretizare: *International Open Data Charter*². Elementul central al noului manifest statuează că „datele deschise sunt un bun public cheie pe care oamenii îl pot utiliza pentru a genera valoare, idei și servicii pentru a crea o lume mai bună pentru toți”.

Uniunea Europeană a preluat principiile deschiderii datelor încă din debutul anilor 2010. Dezbaterile publice și inițiativele de politică publică s-au multiplicat între timp. Formularea curentă este dată de *Open Data Directive*, mai exact de Directiva 2019/1024 a Parlamentului European și a Consiliului din 20 iunie 2019 privind datele deschise și reutilizarea informațiilor din sectorul public.

România s-a aflat în avangarda eforturilor de a deschide date, ediția din 2017 a raportului anual *Open Data Maturity* definind drept *trendsetter* (Carrara et al, 2017). Între timp însă, luând calcul ediția 2020 a aceluiași raport, poziția României a scăzut spre jumătatea clasamentului, dacă nu mai jos (van Knippenberg, 2020b).

Încercăm în volumul de față să descriem impactul portalului guvernamental românesc de date deschise în societatea românească. Evaluarea impactului este însă complicată din rațiuni pe care le explicăm în extenso în capitolul dedicat metodologiei. Pe scurt, ne confruntăm cu dificultăți comune tuturor încercărilor similare din alte țări, date în primul rând de absența unor studii de evaluare a situației inițiale și, mai ales, de surprindere a impactului real al deschiderii datelor. Cel din urmă se manifestă cel puțin la nivel de așteptări în toate sferele vieții publice, incluzând economicul, organizarea administrației publice, în cel social, și în ce privește protecția mediului.

Volumul de față caută să descrie impactul în aceste sfere pe patru direcții majore: notorietatea datelor deschise și a portalului data.gov.ro, legitimitatea acordată deschiderii datelor și portalului guvernamental de actorii cheie, disponibilitatea datelor și utilizarea acestora.

Pentru a înțelege impactul și starea curentă, folosim triangularea informației observată în moduri diferite. Mai exact, au fost realizate 7 studii distincte:

¹ *Who we are (Cine suntem) – platforma Open Data Charter*, <https://opendatacharter.net/who-we-are/>, accesată 7.07.2020.

² *International Open Data Charter*, https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf

- Analiza accesărilor portalului și a designului acestuia
- Studii de caz cu exemple de reutilizare a datelor deschise.
- Analiza opiniilor experților, realizată printr-un studiu de tip Real-time Delphi Panel
- Analiza opiniilor angajaților din administrația publică
- Analiza opiniilor mediului academic
- Producerea unui chestionar de satisfacție a utilizatorilor data.gov.ro și analiza răspunsurilor obținute în prima lună de la implementare
- Analiza comparativă a relațiilor dintre deschiderea datelor și indicatori macroeconomici și sociali.

Cele șapte studii sunt integrate în rezultatele finale, oferind o imagine de ansamblu a poziționării data.gov.ro în spațiul românesc și a impactului acestuia.

Rezultatele relevă prezența unui impact moderat spre mare a data.gov.ro și nevoia de rafinare a implementării. Pentru cea din urmă sunt realizate recomandări la finalul acestui volum.

Pentru fiecare tip de instrument folosit, prezentăm mai întâi premisele utilizării sale și metodologia aferentă, introducem apoi rezultatele, iar secțiunea finală aduce cu sine interpretări în ce privește impactul observat al data.gov.ro și recomandări rezultate din inspectarea rezultatelor. În final, un capitol dedicat triangulează concluziile derivate din analiza datelor colectate prin fiecare din instrumentele utilizate și le rafinează într-o evaluare globală a impactului portalului guvernamental de date deschise. În această secțiune finală, re-utilizăm recomandările menționate în fiecare subcapitol, le îmbogățim, și propunem o listă de acțiuni pe care data.gov.ro le poate întreprinde pentru a crește impactul așteptat.

În interiorul capitolelor, acolo unde a fost posibil, am detaliat discuția despre impact pentru cele patru direcții specifice: notorietatea, legitimitatea, disponibilitatea (datelor), și utilizarea acestora. Analizăm pe rând fiecare aspect, atât din punct de vedere al stării curente, dar mai ales încercând să observăm modificările de stare în timp.

Concluziile relevă valența integratoare a portalului, ca element legitim de accelerare a deschiderii datelor. Arătăm că au fost realizați pași importanți în a crește legitimitatea, observabili mai ales la nivelul administrației publice, dar mai este spațiu de a crește notorietatea. Cele două elemente sunt esențiale în ce privește disponibilitatea și utilizarea datelor. Acestea din urmă sunt în creștere și produc impact, dar e încă un drum lung de parcurs spre a utiliza datele la un nivel mai apropiat de potențialul lor. Utilizarea depinde pe de o parte de potențialul societății în ansamblu de a înțelege, utiliza, reutiliza, și deschide datele. De aici principala recomandare a acestui volum, aceea de a investi în oferirea de exemple sau modele de utilizare. Promovarea unor astfel de modele (realizată deja pe scală largă în țările ce sunt performere europene în deschiderea datelor) constituie cheia efortului de a moderniza și contemporaneiza abordarea datelor și a societății românești în ansamblul său, fiind o obligație implicită a demersului guvernamental, ce derivă din misiunea sa principală, cea de a susține dezvoltarea societății. Includem aici efortul de a sprijini mediul de business, cel ONG, și societatea în ansamblu ei pe calea sa spre dezvoltare.

Dar aceste concluzii sunt detaliate în amănunt în capitolele ce urmează.

OPȚIUNI METODOLOGICE

Metodologia acestui exercițiu de evaluare a impactului pornește de la inventarierea modurilor specifice în care evaluarea impactului deschiderii datelor este realizată în alte exerciții similare, apoi construim propria perspectivă și propunem instrumentele pe care le folosim în restul volumului de față.

EXEMPLE DE ANALIZE DIN ALTE ȚĂRI

Primele evaluări ale mișcării open data au fost simple măsurări ale progresului înregistrat de țările G8 în ce privește îndeplinirea celor 5 principii originale din Charta Open Data (Castro & Korte, 2015). Zuiderwijck & Janssen (2014), în trecerea lor în revistă a modului de evaluare a politicilor open data, notau absența evaluărilor de impact și se centrau exclusiv pe indicatori de performanță, concentrați mai ales pe inputul politicilor de tip open data (e.g. număr de seturi de date disponibile) și mai puțin pe utilizarea acestora (indicator de output). Ei notau însă importanța reacțiilor la politica respectivă (în cazul nostru, legitimitatea) și la nevoile consumatorilor (în cazul nostru acesta se poate converti în cererea de date și în evaluări ale utilizatorilor).

Principalul indicator de input utilizat a fost și încă este bazat pe metricile legate de vizitarea portalurilor de date. (Blank, 2019; Deloitte, 2018; European Commission, 2018; GRU, 2018; van Knippenberg, 2020 etc.). Acesta este de fapt un indicator de vizibilitate sau notorietate a portalurilor, dar spune puțin despre utilizarea lor ca atare.

Un alt mod de a măsura impactul este legat de evaluări ale experților privind impactul deschiderii datelor, așa cum procedează Open Data Maturity Report 2019 (Blank, 2019) sau Raportul DESI (European Commission, 2020).

O metodă mai sofisticată este combinarea celor două acțiuni de mai sus:

- Pe de o parte, așa cum procedează Frank et al (2020): prezența portalurilor și numărul de seturi de date pot fi însoțite în această perspectivă de opinii ale experților privind calitatea, accesibilitatea și relevanța datelor distribuite prin portaluri.
- Pe de altă parte, Open Data Maturity Report 2019 (Blank, 2019) propune o evaluare ceva mai complexă a impactului politicilor OD. Impactul este structurat pe cinci dimensiuni (p.37 și următoarele): monitorizarea permanentă a reutilizării datelor (etichetată drept „strategic awareness”), respectiv reutilizarea în domeniul „politic” (în fapt e vorba de administrația publică și ca posibilitate de control civic asupra acțiunilor decidenților publici și a politicienilor), social, de mediu, și economic. Raportul observă clar noutatea evaluării de impact din domeniu și incoerența acesteia în majoritatea țărilor europene: „[...] *it becomes clear that most countries lack a measurement framework, and impact assessment is unsystematic*” (Blank, 2019:38). De altfel, dincolo de bunele intenții, raportul propune o măsură bazată mai degrabă pe opinii despre impactul posibil, decât pe evaluarea sa ca atare.

Indicatori de input

În lume, evaluarea inițială a impactului a avut în centrul său indicatorii de input (spre exemplu câte seturi de date sunt disponibile), nu pe cei de output (indicatori de utilizare efectivă a seturilor respective). Mai mult, în 2014, Zuiderwijck & Janssen (2014) observau că practic nimeni nu măsurase impactul efectiv al deschiderii datelor.

Pentru a înțelege mecanismele de utilizare a portalurilor, au apărut în ultima vreme analize de varii tipuri. Fără a constitui un corpus extins de literatură, ele aduc în prim plan preocuparea de a înțelege cum funcționează procesul și de a crește impactul fără a-l măsura efectiv.

Rapoarte din Danemarca (Thorup et al, 2019) și portalurile de date din Spania (datos.gob.es/iniciativas) sau Irlanda (<https://data.gov.ie/it/impact>) inventariază exemple de reutilizare a datelor sau de deschidere a acestora, în diferite sectoare. Scopul este de a exemplifica și de a stimula reutilizarea datelor. Astfel de inventarii constituie însă și o modalitate simplă de a exemplifica impactul acțiunilor legate de deschiderea datelor. De altfel, în cazul spaniol și în cel irlandez, pentru a ajunge la „harta inițiativelor” trebuie aleasă secțiunea „impact” în meniul site-ului de date deschise (Impacto/Mapa de iniciativas, respectiv „More/Impact stories”).

Metode de
înțelegere a
factorilor ce
conduc la un
impact mai
ridicat

Studiile de caz aduc informație despre cum funcționează deschiderea datelor în cazuri concrete. Ele pot fi folosite spre exemplu în a studia intermedierea transferului dintre producător și utilizator final (Maail, 2017) sau reutilizarea datelor (Thorup et al, 2019).

Analize mai mult sau mai puțin sofisticate, caută să arate cum deschiderea datelor are impact în economie sau în sectoare ale acesteia (Carpenter & Watts, 2013; Deloitte, 2018; Herala, 2018; PwC, 2017). Majoritatea se rezumă deocamdată la evaluări subiective ale stakeholders-ilor sau experților, iar uneori sunt încercate modele econometrice.

Măsurarea impactului deschiderii datelor și a portalurilor dedicate este o provocare metodologică în sine, iar rapoartele existente se izbesc de nevoia de a analiza un domeniu aproape complet nou, în absența unor modele validate, precum și într-o cultură încă incipientă a proiectelor de evaluare a impactului.

Rămâne un deziderat al cercetării din domeniu să identifice moduri de evaluare a impactului deschiderii datelor în absența unor grupuri de control: măsura fiind practic una globală, nu există (cel puțin momentan!) universuri paralele în care să experimentezi deschiderea sau nu a datelor și să observi care este impactul acesteia ținând sub control toate celelalte variabile.

Suntem prin urmare în situația de a experimenta căutând metodele cele mai potrivite pentru a evalua impactul social, economic, ecologic, administrativ („regulatory”) al măsurilor de deschidere a datelor. Așa cum am explicat, principala problemă cu care ne confruntăm este cea legată de absența grupurilor de control, precum și a măsurătorilor stării inițiale. Cercetarea de față caută să rezolve aceste dificultăți prin apelul la evaluări subiective asupra impactului, comparația cu alte țări, comparația cu sondajul din 2017 realizat cu administrație publică.

PRINCIPIILE EVALUĂRII IMPACTULUI, LA MODUL GENERAL

Pentru a evalua impactul unui proces oarecare, în condiții ideale este nevoie de cinci elemente:

- O evaluare a situației inițiale. Acest lucru presupune dezvoltarea unui set de indicatori de urmărit (Y), a căror schimbare va fi monitorizată permanent, fiind pusă în relație cu derularea procesului studiat.
- Prezența unor indicatori de implementare a procesului (X). În cazul nostru, poate fi vorba de numărul de seturi distribuite prin data.gov.ro, de numărul celor expuși la acțiunile de popularizare, la numărul producătorilor de date ce utilizează data.gov.ro, la gradul de standardizare a seturilor de date, alți indicatori de calitate a datelor șamd.
- Prezența unui “grup de control”, în care nu se petrece procesul studiat.
- Menținerea identică a tuturor parametrilor în grupul de control și în cel în care se aplică procesul.
- O evaluare a situației finale folosind același set de indicatori, Y.

Având aceste patru elemente, este ușor apoi să evaluezi impactul lui X asupra lui Y. Să zicem că situația inițială era dată de un vector Y_0 , iar cea finală de Y_1 . Diferența dintre Y_1 și Y_0 în cazul grupului de control este dată de schimbarea naturală, în absența procesului studiat. Diferența corespunzătoare grupului studiat include schimbarea naturală și impactul datorat procesului. De aceea este esențial grupul de control, pentru a putea elimina schimbarea naturală și să evaluăm impactul pur al procesului.

Cu alte cuvinte, **impactul poate fi definit drept schimbările în situația finală față de situația inițială a obiectului studiat datorită intervenției studiate**. Este esențial prin urmare să cunoști punctul de start și punctul final (sau intermediar, atunci când evaluarea este una la momentul intermediar), să ai un set de indicatori măsurabili, să măsoară expunerea la intervenție, și să ai un grup de control (neexpus la intervenție) sau o expunere diferențiată la intervenția realizată. În cazul nostru intervenția este dată de acțiunilor de deschidere a datelor și de promovare a deschiderii datelor. Sunt mai mulți indicatori de impact ce pot fi urmăriți, descriși în continuarea acestei metodologii și sintetizați în final.

În lumea reală, situația experimentală descrisă mai sus este greu de întâlnit. În implementarea de politici naționale, ba chiar globale cum este deschiderea datelor, grupul de control practic nu există. Atunci se poate experimenta observând intensitatea diferențiată a expunerii la măsurile deschiderii datelor.

Problema metodologiei de față este însă și mai complexă: dispunem de măsuri incomplete ale setului de indicatori definit mai sus drept “X” și lipsește evaluarea situației inițiale (baseline study). Acest lucru nu este neobișnuit în sine. În capitolul anterior am arătat că toate studiile de impact din domeniu se confruntă cu aceeași situație.

În aceste condiții, am dezvoltat un set de principii care să călăuzească acțiunea de măsurare a impactului propusă.

Principiile propuse sunt cele sintetizate în Tabelul 1. Ele statuează mai întâi nevoia de a măsura corect impactul și de a evita erorile de atribuire a schimbărilor generate de alți factori (acuratețea). Așa cum am afirmat în debutul acestei secțiuni, căutăm metode de a compensa absența unui baseline study. În acest sens facem apel la date deja existente ce pot constitui o diagnoză a stării inițiale sau compensăm prin colectarea retroactivă de informație. Dintre metodele propuse în secțiunile următoare, sondajul cu administrația publică reflectă primul tip de abordare, iar cel cu mediul academic pe cel de-al doilea.

Este esențial să colectăm informație și să evaluăm impactul folosind perspective provenite de la toți actorii relevanți implicați în procesul datelor deschise. În una dintre secțiunile următoare reluăm argumentul acesta, arătând că este importantă triangularea informației provenite de la categorii diferite de actori pentru a evita riscul unei evaluări distorsionate a impactului.

Tabelul 1. Principiile metodologiei

- acuratețe
- compensarea preocupării reduse pentru evaluarea impactului din trecut (adică compensarea absenței baseline studies)
- atingerea tuturor categoriilor esențiale de stakeholders
- triangularea metodelor și surselor de date
- completitudine: acoperirea temelor relevante
- sustenabilitate: elaborarea unei metodologii care se poate ușor replica la intervale regulate
- implementare imediată (fezabilitate)
- realizarea la costuri acceptabile (eficiență)
- aplicativitate: capacitatea de a genera recomandări

Triangularea este importantă, ca în orice studiu, și atunci când facem referire la metodele folosite.

Completitudinea este abordată în secțiunea imediat următoare când discutăm la modul general despre temele majore de abordat.

Metodologia pe care o propunem nu este concepută ca un exercițiu de moment,. Evaluarea impactului trebuie să fie permanentă și să poată fi mereu adaptată la evoluții noi. Ceea ce propunem este pretabil la repetare în orice moment, dar recomandăm repetarea la 1-2 ani, pentru a putea lăsa suficient timp în care să se observe schimbări și a nu suprasolicita posibili informatori.

Fezabilitatea este legată de capacitatea de a implementa imediat metodologia propusă, fără a depinde de alte aprobări complicate. Complementar, implementarea trebuie să poată fi realizată la costuri acceptabile.

În fine, scopul final al metodologiei este de a propune un set de metode ce măsoară impactul, al căror aplicare să poată conduce la feedback ce permite ajustarea politicilor de deschidere a datelor în scopul eficientizării acestora și a maximizării impactului.

PRINCIPIILE METODOLOGIEI PROPUSE

Demersul de față ține cont de existența a trei niveluri diferite de diversificare. Pe de o parte e vorba de direcțiile analizei, pe de alta de domeniile în care poate apărea impact, și, în fine, despre actorii care pot furniza informații.

Impactul data.gov.ro se manifestă în cel puțin 4 direcții distincte, care se constituie în întrebări cheie de investigat în acest volum:

- **Notorietatea:** În ce măsură sunt cunoscute cele două obiecte ale studiului – portalul data.gov.ro, respectiv procesul de deschidere a datelor. Notorietatea precede utilizarea pe care o determină (Monino & Sedkaoui, 2016). Cât de mult se știe despre existența datelor deschise și promovarea lor de către este prin urmare întrebarea de început. A nu ști deja aproape te exclude de la posibilitatea unui impact. A ști constituie un prim pas. În consecință, instrumentele pe care le propunem au nevoie să testeze mai întâi notorietatea.
 - **Legitimitatea:** Cât de legitimă este deschiderea datelor este o altă întrebare esențială. Legitimitatea constituie un prim pas spre funcționarea procesului de deschidere. Ea nu este în mod necesar condiționată de notorietate, dar de regulă depinde de aceasta. Legitimitatea acționează ca oportunitate sau obstacol când este vorba despre producătorii de date.
- Identificarea unei legături pozitive între expunerea la acțiunile guvernamentale din domeniul datelor deschise și gradul de legitimitate a procesului poate fi testată prin analize de regresie sau prin analize de conținut în studiile cantitative sau, respectiv, în studiile de caz.
- **Disponibilitatea:** O altă întrebare importantă este dacă a crescut sau nu oferta de date publice, dacă numărul seturilor de date oferite în spațiul public și în particular pe data.gov.ro a crescut sau nu. Ne întrebăm de asemenea ce fel de date sunt disponibile, care este calitatea, formatul, forma acestora.
 - **Utilizarea:** dacă datele sunt folosite și dacă utilizarea (și reutilizarea) conduce la vreun impact

Impactul poate fi identificat într-o serie de domenii ale impactului:

Politic și administrație publică	(a) îmbunătățirea guvernării; (b) întărirea controlului civic; (c) crearea de oportunități de dezvoltare; (d) rezolvarea unor probleme publice.
Economic	(e) ajută agenții economici să își eficientizeze activitatea (f) creștere macroeconomică
Protecția mediului	(g) permite observarea calității mediului și dezvoltarea de acțiuni, programe și politici de mediu
Social	(h) crește participarea civică (i) redistribuire mai bună prin politicile sociale promovate la nivel societal sau local

Bazat pe: Verhulst și Young (2016)

Categoriile de actori cheie de la care pot fi colectate date referitoare la impact:

- Autoritățile publice: sunt deopotrivă utilizatori, intermediari și producători de date (Davies et al, 2013; Maail, 2017; Martin, Turki 2018; Vasiu, 2015)
- Mediul de afaceri: producători, utilizatori. (Vert & Vasiu, 2015)
- Mediul academic: producător, utilizator, uneori intermediar (PwC, 2017; van Knippenberg, 2020)

- ONG: producători, utilizatori, intermediari (Jäppinen et al, 2013; van Schalkwyk et al, 2016)
- Mass-media: producători, utilizatori (Stoneman, 2015)

Cu alte cuvinte, demersul de față se construiește prin a colecta date de la aceste tipuri de actori, pe direcțiile menționate și de a observa potențial impact în domeniile specificate.

METODE DE COLECTARE A DATELOR

Consecința principiilor de mai sus este structurarea unui ansamblu de opt metode diferite pentru colectarea datelor, sintetizate în Tabelul 1. Metodele sunt folosite separat, în analize prezentate pe rând în capitolele dedicate care pot fi regăsite în continuarea volumului de față. La final, în capitolul de concluzii, utilizăm rezultatele aplicării fiecărei metode pentru a structura o privire de ansamblu asupra impactului deschiderii datelor și a portalului data.gov.ro.

Tabelul 2. Sinteza metodelor folosite pentru colectarea datelor

acțiune	scop	Observații
Back-log analysis a data.gov.ro	Analiza utilizării primare	În limita posibilităților tehnice ale data.gov.ro Nu indică în mod direct impactul
Dezvoltarea unui chestionar de feedback asociat data.gov.ro	Analiza potențialului de reutilizare	Permite dezvoltarea unui mecanism de evaluare permanentă a impactului și de ajustare din mers a portalului
Analiza comparativă a câtorva portalurilor guvernamentale europene de date deschise	Comparația între România și alte țări	Permite evaluarea poziționării României în comparație cu alte țări și creșterea eficienței data.gov.ro prin împrumut de idei.
Studii de caz cu utilizatori corporativi și non-utilizatori corporativi ai data.gov.ro	Ilustrarea modului în care se produce impactul	Nu măsoară impactul în sine, ci doar îl ilustrează. Metoda este folosită de multe rapoarte pentru a suplini absența unor măsurători mai precise
Sondaj/Panel Delphi cu experți	Rafinarea impactului observat	Compensează lipsa de date privind starea inițială
Sondaj cu administrația publică	Notorietate, legitimitate, utilizare, reutilizare, oferire de date, impact în varii domenii	Comparația cu 2017 asigură evaluarea consistentă a impactului
Sondaj cu mediul academic	Notorietate, legitimitate, utilizare, reutilizare, oferire de date, impact în varii domenii	Evaluarea impactului se bazează pe relatările celor chestionați cu privire la situația lor în urmă cu câțiva ani
Analiza comparativă	Evaluare globală a impactului	Impactul se referă la omenire în modul general, nu este particularizat pe România, dar se aplică și acesteia

Pentru a ușura lectura, furnizăm specificații pentru fiecare metodă în parte în capitolul dedicat metodei respective. În continuarea acestei secțiuni explicăm însă pe scurt rațiunile utilizării fiecărei metode, reiterând astfel cadrul general în care se plasează fiecare demers de cercetare în parte și modul în care ele servesc scopului comun al acestui volum.

Analiza accesărilor data.gov.ro (back-log analysis) ne oferă o primă informație despre utilizarea de date deschise. Fără îndoială informațiile sunt parțiale: ele se referă *doar* la data.gov.ro și indică doar accesări, nu și utilizări ale datelor. Totuși, astfel de informații, așa cum am explicat, fie ele și limitate, sunt cele care permit o primă și rapidă luare de contact cu scopul nostru: analiza impactului în general și, în particular, observarea utilizării datelor.

Dezvoltarea chestionarului de feedback constituie tot un element legat de utilizarea portalului data.gov.ro. Chestionarul a fost implementat pe site-ul data.gov.ro, apoi am analizat datele colectate timp de trei luni (decembrie 2020-martie 2021). Scopul a fost să înțelegem comportamentul utilizatorilor.

Analiza unui număr de portaluri europene de date deschise a permis înțelegerea modului în care acestea funcționează în țările care înregistrează cele mai mari scoruri ale deschiderii datelor și în țările similare României.

Studiile de caz cu potențiali utilizatori corporativi au avut menirea de a evidenția notorietatea, utilizarea, reutilizarea, dar și legitimitatea deschiderii datelor și disponibilitatea datelor. Au fost selectate companii care utilizează data.gov.ro și companii care ar putea utiliza dar nu o fac.

Rațiunile selectării metodei sunt destul de simple. Exemple de utilizare în mediul privat, societate civilă sau administrație publică sunt furnizate în literatura existentă ca modalități de a ilustra impactul datelor deschise. Raportul Open Data Maturity 2019 (Blank, 2019) procedează astfel pentru a compensa lipsa de date în ce privește impactul social (pag. 41-44). Cu același scop, și prezenta metodologie propune prezentarea a 4-6 studii de caz în care datele deschise ilustrează impactul acțiunilor guvernamentale în domeniul OD în domeniul social, al administrației publice, al protejării mediului, etc. Au fost incluse și exemple de utilizare *not-for-profit*, așa cum sunt cele care vizează transparența sectorului publice sau promovarea punctelor de interes cultural. Analiza în fiecare studiu de caz a avut în prim plan grila de citire sugerată de Verhulst și Young (2016), căutând să identifice contribuții în cele patru domenii: *improving government; empowering citizens; creating opportunity; solving public problems*. Deși, așa cum am argumentat deja, exemplificarea prin studii de caz este prezentată și în alte rapoarte ca o analiză sui-generis de impact (a se vedea spre exemplu Blank, 2019; Maail, 2017; sau datos.gob.es/iniciativas), opțiunea este însoțită de un risc ridicat de a evalua eronat impactul real. Simpla selectare a studiului de caz face ca rezultatele sale să fie simple observații particulare legate de caz oarecare. Chiar și când sunt mai multe studii de caz, dacă nu devin repetitive, ele nu ilustrează impactul în întreaga sa manifestare. Studiile de caz au însă proprietatea deosebită de a genera ipoteze rapide despre factorii care conduc la un impact ridicat și, în plus, promovează programele de open data, catalizând impactul.

Panelul Dephi cu experți din domeniu a fost realizat online, având scopul de a plasa opinia experților cu privire la oportunitățile, legitimitatea, provocările deschiderii datelor și cu privire la prezența data.gov.ro. Am ținut cont de premisa deja menționat mai sus că experții sunt cei care pot avea o reprezentare consistentă despre utilizarea și beneficiile datelor deschise. Aceasta este calea urmată de Open Data Maturity Questionnaire (Blank, 2019). Am construit o cercetare de tip real-time Delphi (realizată online prin soluția furnizată de Kialo).

Rolul esențial al administrației publice ca utilizator, producător, intermediar, reglementator, și catalizator al deschiderii datelor face esențială investigarea ei ca agent cheie de generare a legitimității și notorietății. A observa care este atitudinea administrației publice față de deschiderea datelor și reutilizarea lor este extrem de important pentru a înțelege contextul acțiunilor guvernamentale în domeniul datelor deschise.

În legătură cu această categorie de actori ne-am aflat în situația privilegiată de a dispune de un sondaj de opinie realizat în rândul funcționarilor publici în 2017, privind atitudinea față de datele deschise (Voicu, 2018). Am putut prin urmare să repetăm sondajul, și să comparăm cifrele de atunci cu cele din prezent. Mai mult decât atât, am avut opțiunea de a derula sondajul ca panel: aproximativ 40% din cei aproape 1000 de respondenți din 2017 au indicat disponibilitatea de a fi recontactați și putem pune în legătură opiniile lor din trecut cu cele curente, precum și comportamentele legate de datele deschise. Controlând pentru expunerea la acțiunile guvernamentale în domeniul datelor deschise am putut evalua impactul efectiv în ce privește notorietatea, cunoașterea, legitimitatea. De asemenea, au fost estimate frecvența comportamentelor de reutilizare, diseminare și intermediere a transiterii datelor.

Rațiunea unui sondaj similar adresat mediului academic a fost deja argumentată în secțiunea privind actorii cheie ai procesului deschiderii de date. Am arătat acolo că metoda este deja folosită în alte societăți. Opțiunea aleasă a fost cea a unui websurvey, cu ponderare ulterioară pentru a asigura un eșantion calibrat conform caracteristicilor populației de referință. Temele din chestionar sunt similare celor din chestionarul cu administrația publică, pentru a asigura comparabilitatea. Se adaugă unele specifice mediului academic, inclusiv itemi comuni cu cei raportați de PwC (2017), pentru comparații cu alte societăți. Replicabilitatea (anuală/bianuală) este similară sondajului cu funcționari publici.

Analiza comparativă prezentată este una scurtă, dar care are natura de a exemplifica efectele deschiderii datelor. Comparația cu alte țări permite practic evaluarea impactului. Am folosit GDP/capita ca indicator de output, prezis cu Open Data Maturity index (ODMI) în modele de tip fixed effects (modele panel), evaluând cu cât contribuie schimbarea în deschiderea datelor la dinamica economică a țărilor europene. Impactul observat al ODMI constituie un impact așteptat mediu. El nu se referă la România, dar furnizează informație privind evoluțiile așteptate ale României în funcție de deschiderea datelor. Cauzalitatea este asigurată prin succesiunea cronologică a măsurătorii, dar se impune o precauție importantă: impactul estimat este bazat pe experiența trecută. Este posibil ca la grade diferite de deschidere a datelor relațiile observate, impactul să se modifice.

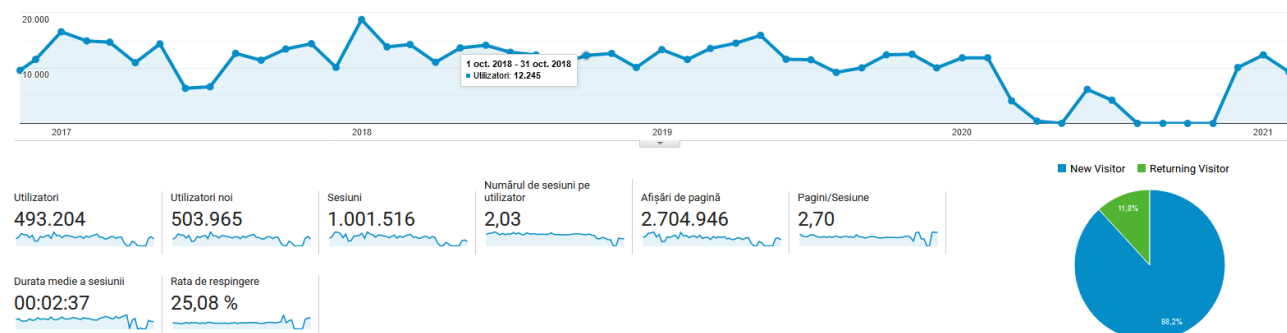
BACK-LOG ANALYSIS

Data.gov.ro folosește exclusiv Google Analytics pentru a contabiliza accesările site-ului. Abia din decembrie a fost adăugat un chestionar de satisfacție. Lipsesc însă statistici privind numărul de descărcări ale datelor propuse.

Data.gov.ro include în prezent (24.02.2021) 2.182 seturi de date, structurate pe furnizori, domenii largi de care aparțin, tipuri de formate etc. Așa cum spuneam, nu putem ști exact câte din aceste seturi sunt în fapt utilizate și nici despre gradul lor de atractivitate în timp.

Putem în schimb observa, folosind Google Analytics, cum a variat accesarea site-ului în timp. Figura 1 ilustrează evoluția din noiembrie 2016 (primele înregistrări disponibile în ce privește numărul de utilizatori) și până în prezent. Conform Google Analytics este vorba de aproximativ 500.000 de utilizatori unici ce au accesat data.gov.ro în această perioadă. Trendul observat este unul staționar: practic avem cam același număr de utilizatori lunari din 2016 până în prezent. Lunile unde se observă scăderi sunt de regulă asociate perioadelor când scriptul de la Google Analytics implementat în data.gov.ro nu a funcționat din varii motive. Întotdeauna luna februarie are mai puțini utilizatori, dat fiind că este mai scurtă cu 1-3 zile decât celelalte luni. În plus, statistica nu acoperă complet februarie 2021 care nu se încheiase la momentul extragerii graficului.

Figura 1. Utilizatori pe lună noiembrie 2016 – februarie 2021



* În lunile cu trafic evident mai mic, setările Google Analytics pe data.gov.ro nu au funcționat. Cifra pentru ultima înregistrare (februarie 2021) este firesc să fie mai mica, dat fiind că nu acoperă întreaga lună.

Figura 2. Durata medie a sesiunii: februarie 2015 – ianuarie 2021, medii pe zile



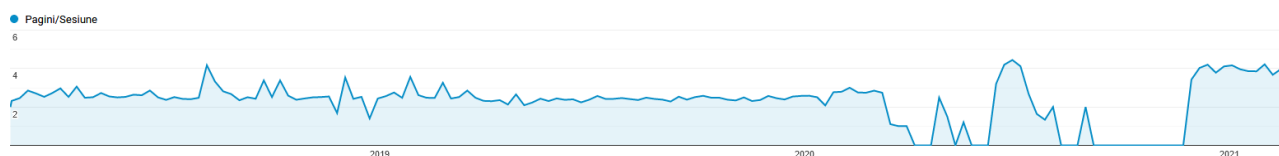
* În lunile cu trafic evident mai mic, setările Google Analytics pe data.gov.ro nu au funcționat.

Durata medie petrecută pe site este una redusă, de circa 2 minute. Cum a selecta datele dorite și a le descărca durează probabil cel puțin 5 minute, avem o indicație asupra faptului că cel mai probabil doar o zecime dintre utilizatori au descărcat efectiv date de pe portal. În plus, observăm că, după începerea anului 2016, apare un trend de creștere lină dar stabilă a timpului petrecut pe site. Aceasta se întrerupe în momentul când Google Analytics nu mai funcționează, care coincide cu perioada pandemică curentă. Dacă luăm în considerare doar

trendul anterior lui 2020, creșterea este evidentă. Mai mult, în decembrie 2020-ianuarie 2021, când raportarea Google Analytics nu este afectată de salturi în funcționare, numărul mediu de minute petrecute pe site revine la normal, încadrându-se în trendul ascendent remarcat anterior lui 2020.

Cu precauția absenței datelor pentru 2020-2021, Figura 3 evidențiază faptul că accesarea până în 2020 s-a plasat în jurul unei medii de 2-3 pagini pe sesiune, dar a crescut în 2020 la 4 pagini. Cum viteza de accesare a site-ului nu a scăzut, observația constituie un indiciu dublu. Pe de o parte este vorba despre creșterea complexității site-ului sub aspectul numărului de seturi de date disponibile. Pe de altă parte, coroborat cu timpul mai mare petrecut pe site, este vorba probabil de o creștere a interesului utilizatorilor.

Figura 3. Pagini pe sesiune, medii săptămânale, februarie 2018-februarie 2020



* În lunile cu trafic evident mai mic, setările Google Analytics pe data.gov.ro nu au funcționat.

Tabelul 3. Statistici pentru perioada februarie 2018 – februarie 2021

	Utilizatori	Utilizatori noi	Sesiuni	Pagini/Sesiune	Durata medie a sesiunii
Total	320.590	324.311	618.562	2,66	00:02:43
<i>desktop</i>	216.698(67,80 %)	220.976(68,14 %)	499.813(80,80 %)	2,55	00:03:03
<i>mobile</i>	97.176(30,41 %)	97.607(30,10 %)	112.051(18,11 %)	3,08	00:01:18
<i>tablet</i>	5.727(1,79 %)	5.728(1,77 %)	6.698(1,08 %)	3,66	00:02:02

Aceste creșteri, chiar dacă nu sunt impresionate, devin mai importante ca dimensiune dacă ținem cont și de faptul că, în această perioadă, au apărut accesările și prin portalul european de date, ceea ce a diminuat probabil o parte a traficului potențial.

STRUCTURA SITE-ULUI DATA.GOV.RO

Pentru a înțelege mai bine structura data.gov.ro, am analizat site-urile a opt alte portaluri europene similare. Analiza nu are în vedere cantitatea sau calitatea datelor, ci nivelul de autopromovare a site-ului, gradul în care acesta este prietenos cu utilizatorii.

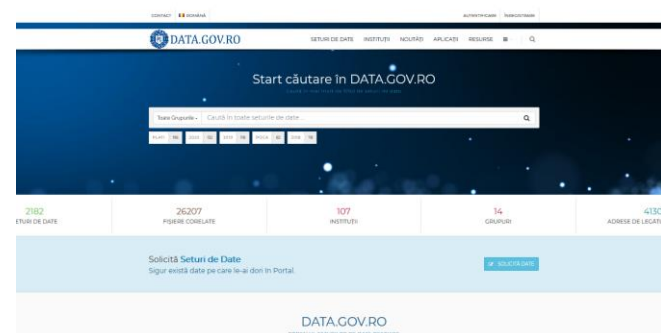
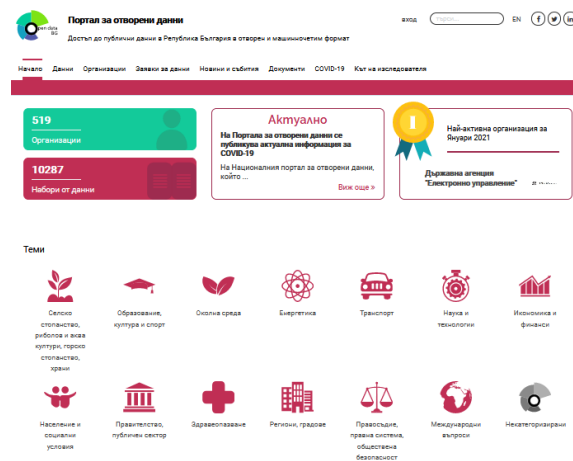
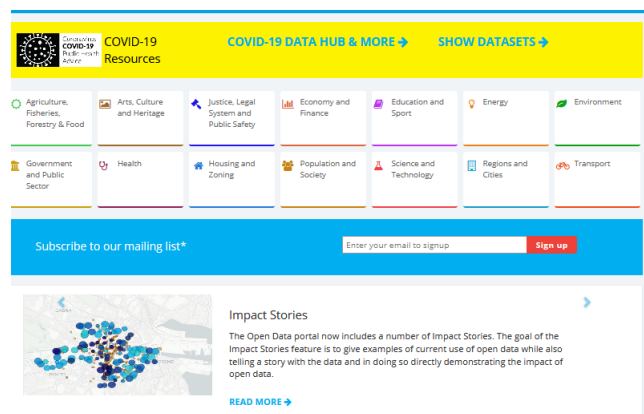
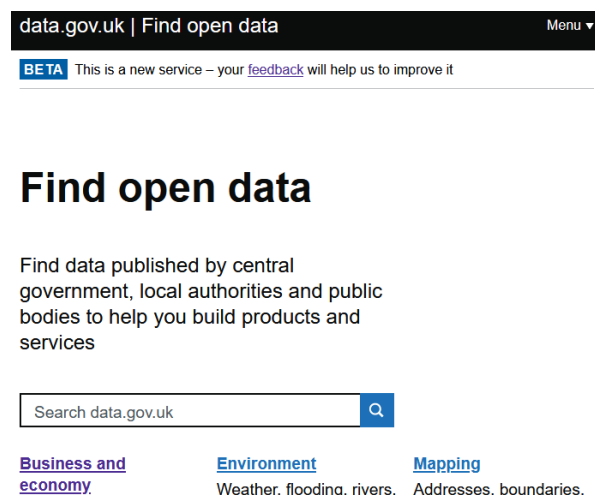
Criteriile de analiză luate în calcul sunt legate de formatul site-ului (dacă este dinamic, dacă include o secțiune vizibilă de știri, dacă este bogat în imagini), de modul de structurare a informației (dacă include clasificarea datelor pe domenii, facilitând astfel căutarea și accesarea), dacă include date despre utilizare (altmetrics și posibilitatea utilizatorilor de a lăsa feedback), dacă furnizează exemple de utilizare (criteriu denumit generic „impact stories”, și care indică dacă utilizatorilor li se oferă pattern-uri de urmat pentru a cataliza utilizarea datelor), și dacă site-ul este adaptat la realitatea imediată, prin a avea o secțiune dedicată COVID-19. În plus, am adăugat disponibilitatea informației în alte limbi.

Tabelul 4. Trecere în revistă a câtorva caracteristici ale site-urilor portalurilor guvernamentale de date deschise din câteva țări europene

Site	News	Dinamic	Imagine	Structurare a datelor pe domenii	Altmetrics	Chestionar satisfacție / chat sau forum	Site/secțiune dedicate COVID-19	Impact stories	Limbi
Franța data.gouv.fr	Da	Da	Da, cu multe imagini atractive pentru vizitatori	Da, cu structurare ierarhică	Da, număr de descărcări	Doar comentarii ce pot fi formulate pe site	Da	Da, încă din prima pagină	EN, FR, ES
Irlanda data.gov.ie	Da	Da	Da, cu atenție acordată culorilor și atractivității	Da, extins	Nu, doar openness rating și Google analytics	Nu / da	Da	Da	EN, Gaelic
Danemarca opendata.dk	Da	Nu	Puține, neprietenos	Da	Nu	Nu / nu	Nu	Da	DK
Bulgaria data.egov.bg	Da	Nu	Nu	Da	Nu, doar Google Analytics	Nu / nu	Da	Nu	BG, EN
Ungaria opendata.hu	Nu	Nu	Nu	Da, sumară	Nu	Nu / nu	Nu	Nu	HU (cu unele elemente doar în EN!!?)
Polonia dane.gov.pl	Da, dar în josul paginii	Nu	Nu	Da	Nu, doar openness rating și Google analytics	Nu / nu	Da, un API	Nu, doar API	PL, EN
UK data.gov.uk (versiune beta!)	Nu	Nu	Nu	Da	Nu, doar Google Analytics	Da / nu	Nu	Nu	EN
Spania datos.gob.es	Da	Da	Da	Da, extins și multidimensional	Nu, doar Google Analytics	Nu / da	Nu	Da, extins	ES, EN, Catalană, Bască, Galiciană
România data.gov.ro	Da, dar neprietenos	Nu	Nu	Da	Nu, doar Google analytics	Da / da	Nu	Câteva exemple	RO, EN, IT, FR, ES, DE

Portalurile au fost accesat pe 24.02.2021.

Figura 4. Capturi de ecran a paginii de intrare din câteva portaluri



*de la stânga la dreapta, pe primul rând: ES, UK, FR, iar pe al doilea: IE, BG, RO. Accesare pe 24.02.2021

Selecția țărilor luate în considerare spre comparație a fost realizată astfel: am luat în considerare Franța, Irlanda și Danemarca - primele trei țări în ierarhia evaluării portalurilor din Open Data Maturity Report 2020, care pot fi considerate drept trend-setters. Am adăugat țările din proximitatea geografică a României (Bulgaria, Ungaria și Polonia). Am adăugat Franța, Marea Britanie și Spania, ale căror portaluri atrag cele mai multe citări (a se vedea Tabelul 6 de la pagina 37 a acestui volum). Cum Franța aparține pentru două din aceste categorii, au rezultat opt studii de caz. Opțiunea pentru țările cu cele mai multe citări este dată de nevoia de a avea exemplul portalurilor care au gradul cel mai ridicat de utilizare, chiar dacă suntem conștienți că numărul mare de citări depinde și de dimensiunea țării și nivelul de internaționalizare a limbii vorbite în țara respectivă.

Practicile observate sunt diverse, după cum o relevă Tabelul 4. Portalul maghiar de date practic nu include date. Cel danez oferă o multitudine de căi de căutare, include statistici referitoare la bazele de date aduse în spațiul public, precum și link-uri către portaluri de date deschise ale administrației locale. Portalul irlandez este entuziasmant ca design și nivel de actualizare, cu un flux remarcabil de informații și detalii, oferind o varietate de resurse. Portalul francez și cel irlandez preiau menționări ale lor sau datelor promovate în media și social media. Portalurile danez, bulgar, maghiar sunt statice, nu oferă nimic atractiv privirii, culorile sunt terne, iar designul lipsit de atractivitate și imaginație. Portalul spaniol este extrem de bogat în informație, cu o structurare a seturilor de date pe criterii multiple, dar nu include o secțiune dedicată Covid-19.

Importante sunt însă comparațiile cu portalul românesc de date.

Data.gov.ro se încadrează în rândul portalurilor nu tocmai atractive pentru utilizator, ci doar centrate pe furnizarea de date. Fără a fi la fel de rigid ca și portalul britanic (care pare a fi totuși într-o versiune de tranziție), portalul românesc este lipsit de imagini care să rețină vizitatorul și să îi atragă atenția, de elemente dinamice și de un flux de știri care să lase senzația unui zone eferescente. Spaniolii, irlandezii sau polonezii excelează în acest sens, portalurile lor aducând frecvent știri despre activități desfășurate, despre date noi, despre activități ale altora. Portalul francez preia fluxuri din social media (Twitter). Mici astfel de elemente de marketing permit promovarea portalurilor.

Din punct de vedere al modului de structurare a tag-urilor, portalul românesc este în linia generală a celorlalte site-uri analizate.

Ca deschidere spre utilizator, prezența Google Analytics este în linia generală, dar așa cum am arătat în capitolul precedent, este loc pentru deschidere către altmetrics, așa cum site-ul irlandez deja o face.

Capacitatea curentă de a atrage feedback plasează data.gov.ro ca exemplu pentru restul site-urilor similare.

Majoritatea portalurilor analizate au dedicat secțiuni COVID-19, adaptându-se astfel actualității pe care o trăim și devenind interesante pentru utilizatori. Simpla vizită după căutarea pentru date legate de COVID-19 crește potențialul de cunoaștere a portalului guvernamental și aduce potențiali utilizatori pentru toate celelalte date disponibile pe portal. Data.gov.ro nu include astfel de informații, dată fiind și absența acestora în format structurat în spațiul românesc. Ar putea însă furniza o colecție de link-uri către site-urile ce prezintă informație relevantă despre COVID-19, o practică pe care am observat-o și în alte portaluri.

Majoritatea site-urilor vest-europene acordă o importanță accentuată exemplelor de urmat (coloana „Impact stories” din Tabelul 4). Ele constituie un motor de promovare și repromovare continuu, oferind deopotrivă furnizorilor și utilizatorilor potențiali de date exemple de urmat. Portalul românesc suferă din acest punct de vedere, deși este peste nivelul țărilor din jur.

În fine, data.gov.ro este superior tuturor site-urilor analizate prin deschiderea internațională, fiind prezentat (cel puțin parțial) într-o varietate de limbi.

Concluzia simplă este că portalul românesc este din punct de vedere al designului site-ului peste cel al țărilor din jur, dar destul de departe de performenții Europeni. Data.gov.ro lasă impresia de a fi centrat doar pe furnizarea de date, neacordând aproape de loc atenție impulsiei utilizării acestora sau atragerii și retenției utilizatorilor potențiali. Acestea sunt însă lucruri simple ce pot fi rezolvate prin aducerea designului la un standard mai dinamic, integrând fluxuri de știri, anunțuri despre noutăți și evenimente prezentate grafic mai atractiv, și promovarea exemplelor de utilizare și reutilizare. Integrarea cel puțin a statisticilor privind numărul de descărcări este de asemenea un element ce poate promova site-ul, dar și oferi elemente de evaluare a impactului său.

Studiile de caz au rolul de a ilustra modul de utilizare și reutilizare a datelor deschise prin data.gov.ro. Am ales trei tipuri de studii de caz pe care să le realizăm. Un prim set de entități interesante din perspectiva acestui studiu sunt furnizorii de servicii online care preiau dat direct din data.gov.ro și erau deja listați pe portal ca atare. Am contactat 6 astfel de furnizori, primind acceptul de a discuta de la 4 dintre ei.

Al doilea tip de studii de caz interesante e legată de proiecte care nu folosesc date furnizate de pe data.gov.ro, deși teoretic ar putea-o face. Important în acest context este notorietatea data.gov.ro și tipul de date ce ar putea fi de interes pentru acești utilizatori. Dat fiind rolul mai degrabă de grup de control pentru acest tip de studiu de caz, am realizat unul singur.

Al treilea tip de utilizator este cel academic. Am realizat în acest sens o trecere în revistă a șase articole academice publicate peer-reviewed și care utilizează date de pe data.gov.ro.

Pentru a menține anonimitatea celor ce au discutat cu noi, am denumit cu acronime entitățile respective.

Tabelul 5. Lista studiilor de caz

Categoria		Tip produs	Public țintă	Formă de organizare	Mod de utilizare	Caracterul reutilizării
Reutilizator	R1	Location intelligence	Public larg, mediul de afaceri	Business	Prin site	Public Servicii gratuite și variantă plătită
	R2	Accountability	Mass-media, ONG, mediul politic	ONG	Prin site	Public Servicii gratuite
	R3	Company reputation	Mediul de afaceri	Business	Site, după logare	Servicii cu plată
	R4	Orientare școlară	Public larg (elevi, părinți)	ONG	site	Servicii gratuite
Non-utilizator	U1	Audit financiar	Mediul de afaceri	Business	Particularizat pentru fiecare client	Servicii cu plată
Academic	A1-A11	Cercetare academică	academic	academic	Publicații științifice	Disponibil gratuit sau cu plată

R1: SERVICII DE LOCATION INTELLIGENCE

I. CE FACE R1

R1 este o echipă de lucrători în industria IT, care și-au unit eforturile pentru a furniza un serviciu online de *location intelligence*: pentru un loc selectat de către beneficiar, oferă informații despre zona din jur.

Produsul este în dezvoltare, cu resurse interne. Încă nu este sustenabil, nici nu are o finanțare externă. Echipa alocă un timp limitat dezvoltării, în paralele cu alte proiecte.

Clienții potențiali sunt atât persoane (spre exemplu, persoane care vor să se mute), cât și mici afaceri (care vor să deschidă într-o anumită zonă).

II. DATE UTILIZATE ȘI PRODUSE

Proiectul folosește exclusiv date din surse publice, aproximativ jumătate dintre ele provenind pe de data.gov.ro. Alte surse de date sunt: Ministerul Educației (SIIIR), Institutul Național de Statistică sau Autoritatea Electorală Permanentă. Chiar dacă unele dintre aceste instituții (MEN, AEP) publică date și pe data.gov.ro, seturile publicate nu sunt complete și nici publicate la timp.

Pe termen mediu, există intenția de a prelua date și de la autorități publice locale sau deconcentrate, în special autorizații sanitar-veterinare și autorizații de construire. Acestea însă nu sunt centralizate și apare problema căutării lor pe fiecare pagină în parte.

În toate situațiile, echipa R1 descarcă seturile de date (le duplică pe serverele proprii), le curăță, face geocodare (unde este cazul) și caută/definesc variabilele care unesc datele. Produsul final este un nou set de date, pe care nu l-au deschis. Cei din R1 încă nu au luat în calcul deschiderea datelor produse.

Au în vedere programe de cooperare cu mediul academic. Au avut o astfel de colaborare cu o universitate de stat, unde au pus, gratuit, la dispoziția cercetătorilor date pentru un studiu despre accesul cetățenilor la spații verzi.

III. RELAȚIA CU DATA.GOV.RO

Data.gov.ro este cel mai important furnizor de date pentru R1. De pe portalul național de date deschise preia informații despre: transport, sănătate, afaceri (Registrul Comerțului) și educație (datele despre educație sunt suplimentate cu seturi de date obținute de la Ministerul Educației).

Ministerul Educației publică doar parțial datele pe care le are pe data.gov.ro. Restul datelor sunt extrase din pagina SIIIR cu ajutorul unui script. Similar, AEP publică doar parțial datele. INS nu publică deloc date pe data.gov.ro. Niciuna dintre aceste instituții nu oferă acces la date prin API.

R1 nu folosește API pus la dispoziție de către data.gov.ro. Motivele principale invocate sunt trei:

- Nu există un API unic pentru tot portalul, ci câte unul pentru fiecare set de date;
- R1 a remarcat probleme de stabilitate și funcționare ale data.gov.ro, probleme care au condus la neîncredere;
- Seturile de date folosite se actualizează rar (6 luni, 1 an) și atunci descărcarea lor nu constituie o dificultate.

Datele din data.gov.ro sunt descărcate în propriile servere și, în principiu, sunt îmbunătățite (curățare, geocodare, legare cu alte seturi de date). Datele îmbunătățite nu se întorc în portalul data.gov.ro, ci sunt utilizate exclusiv în aplicația proprie a R1.

Principalele probleme legate de seturile de date din data.gov.ro sunt, în opinia celor de la R1:

- Completitudine: sunt instituții (ex.: Registrul Comerțului) care publică doar parțial date, restul sunt cu plată; sunt instituții (MEN, AEP) care publică doar parțial date, restul fiind pe pagina proprie;
- Acces (vezi discuția despre API);
- Folosirea (in)coerentă a variabilelor de unificare a seturilor de date: SIRUTA, SIIIR, CIF.

IV. VALOAREA ADĂUGATĂ A PROIECTULUI

R1 are potențialul de a transforma seturile de date într-un serviciu nou - location intelligence - util pentru persoane și firme. Pentru că produsul este în dezvoltare, nu există încă date privind impactul său.

R2 ȘI UTILIZAREA BANILOR PUBLICI

I. CE FACE R2

R2 este un think-tank, care a dezvoltat două proiecte de reutilizare a datelor deschise. R2.1 este proiect care urmărește să crească transparența achizițiilor publice, rezultatele sale fiind prezentate prin intermediul paginii de Facebook cu același nume. R2.1 este un proiect non-profit.

Aceeași echipă de voluntari a dezvoltat un alt proiect (R2.2), care își propune să prezinte într-o formă mai ușor de înțeles cheltuielile autorităților publice locale. R2.2 a fost dezvoltat în cadrul unui partid politic.

Cele două proiecte (R2.1 și R2.2) se adresează publicului larg și mass-media.

II. DATE UTILIZATE ȘI PRODUSE

R2.1 folosește datele din portalul național de achiziții publice. Inițial, a folosit datele publicate de data.gov.ro. Ulterior, echipa a decis să preia date direct din portalul de achiziții publice, chiar dacă a însemnat un efort suplimentar de programare. Motivația declarată este legată de slabă calitate a serviciilor oferite de data.gov.ro .

R2.2 a acționat ca model pentru dezvoltarea R2.1. R2.2 folosește date din pagina Ministerului Finanțelor Publice, transparenta.bugetara.gov.ro. Aceste date, deși sunt (relativ) deschise, nu au fost niciodată publicate pe data.gov.ro.

Ambele proiecte oferă două tipuri de produse:

- Vizualizări pentru publicul larg;
- Analiză, infografice și scurte sinteze pentru mass-media

Punctual, la cererea unor jurnaliști sau a unor cercetători, echipa implicată în R2.1 a realizat și prelucrări specifice, în regim de voluntariat.

III. RELAȚIA CU DATA.GOV.RO

R2.1 a folosit inițial datele din portalul de achiziții publice prin intermediul data.gov.ro. A întâmpinat însă mai multe probleme:

- Datele sunt de o calitate scăzută și cer un efort de curățare. Conform celor de la R1, în ciuda atenționărilor, nu au fost făcute niciun fel de demersuri pentru creșterea calității datelor.
- Datele sunt publicate cu o întârziere de minim trei luni. Mai ales pentru mass-media, este un interval nerezonabil de timp.
- Datele sunt incomplete: nu includ toate câmpurile disponibile.
- API-ul pus la dispoziție de data.gov.ro este neprietenos.

În acest moment, R2.1 nu mai folosește data.gov.ro, ci preia date direct din portalul de achiziții publice.

R2.2 nu folosește data.gov.ro, pentru că nu include datele ANAF. Ar prefera să folosească portalul de date deschise, pentru că cel al ANAF este neprietenos.

În ambele situații, echipa ar prefera să folosească data.gov.ro.

IV. VALOAREA ADĂUGATĂ A PROIECTULUI

R2.1 are o audiență de aproximativ 10.000 de persoane, în general persoane informate care pot la rândul lor să distribuie informația. Cele mai populare postări au până la câteva sute de mii de vizualizări și sute de redistribuiri, dar acestea sunt destul de rare.

Informațiile publicate au fost folosite în repetate rânduri de mass-media, mai ales când a fost vorba de achiziții greu de justificat.

Nu sunt disponibile informații despre audiența R2.2.

R3 – BONITATEA COMPANIILOR

I. CE FACE R3

R3 este un serviciu de verificare a bonității partenerilor de afaceri. Folosește date din surse deschise pentru a genera un profil de încredere al firmelor din România. Oferă două tipuri de servicii cu plată:

- Verificarea unei firme (prin abonament care include un număr de verificări)
- Monitorizarea unei firme (prin notificări cu privire la schimbările survenite)

Clienții sunt firme care fac afaceri cu plata la termen și doresc să verifice bonitatea potențialilor parteneri de afaceri.

Produsul are trei ani și este deja profitabil. Este dezvoltat integral cu resurse proprii.

II. DATE UTILIZATE ȘI PRODUSE

R3 folosește date din mai multe surse:

- Data.gov.ro: registrul comerțului, achiziții publice, bilanțuri firme, date privind profesiile liberele.
- ANAF: date despre firme (prin API)
- INS: date statistice generale
- Alte surse nedezvăluite (secret comercial)

După curățarea și concatenarea datelor din alte surse, rezultă un set îmbunătățit de date, dar acesta nu este deschis din rațiuni comerciale. Sunt conștienți că există o concurență puternică pe acest segment de piață și nu doresc să ajute concurenții.

Pentru extinderea portofoliului de servicii, au identificat și alte date publice valoroase, dar acestea nu sunt disponibile sau au un cost descurajator:

- ANCPI: datele privind proprietatea;
- Autorizații de transport;
- ANRE: date privind certificatele energetice;
- Monitorul Oficial: informații într-un format citibil, dacă nu prelucrabil automat;
- AEGRM: arhiva electronică a garanțiilor reale mobiliare;
- ONRC: datele publicate la acest moment sunt minimale.

Și-ar dori ca măcar o parte dintre aceste seturi de date să fie prezente pe data.gov.ro.

III. RELAȚIA CU DATA.GOV.RO

Seturile de date din portalul național de date deschise reprezintă un procent însemnat, aproape jumătate, din datele utilizate de R3.

În general, experiența utilizării data.gov.ro este una rezonabilă, lucru determinat și de faptul că datele folosite sunt unele cu actualizare rară.

Datele sunt descărcate și duplicate pe serverele proprii. Accesul prin API li se pare instabil și nu inspiră încredere.

Datele trebuie curățate, calitatea lor este una medie. Nu au făcut sugestii pentru îmbunătățire, pentru că au văzut deja comentarii similare de la alte persoane.

Principala așteptare de la data.gov.ro este să crească numărul de seturi de date furnizate, în special cu cele enumerate la punctul anterior.

De asemenea, așteaptă ca furnizorii să ofere în format deschis date complete, nu doar o parte a lor. Faptul că numeroase instituții publice cer taxe foarte mari pentru acces la date limitează creșterea afacerii.

Echipa proiectului caută activ idei noi, inclusiv prin urmărirea portalului de date deschise. Publicarea unor seturi de date noi poate genera inovare.

IV. VALOAREA ADĂUGATĂ A PROIECTULUI

R3 este o afacere pe profit după trei ani de activitate. A creat cinci locuri de muncă și contribuie la stabilitatea financiară a clienților (numărul de abonați nu este public).

Afacerea are potențial de creștere, iar acest lucru poate fi facilitat printr-o ofertă mai variată de seturi de date.

R4 – DESPRE MESERII ȘI ORIENTARE ȘCOLARĂ

I. CE FACE R4

R4 este un proiect non-profit, implementat de o organizație civică ce și-a propus să dezvolte soluții și proiecte digitale gratuite de utilitate publică.

R4 este un serviciu de *location intelligence* adresat în principal elevilor de clasa a opta și părinților ce își propun să aleagă o școală. Pentru fiecare loc selectat, oferă informații despre meseriile pregătite în școli din apropiere, și potențialul pieței de forță de muncă din zonă (nu oferă efectiv informații despre locurile de muncă, ci despre potențialul de absorbție al pieței).

Produsul este în dezvoltare, cu resurse interne. Este o inițiativă non-profit. Are potențial de monetizare pe termen lung, prin servicii adăugate pentru firme, dar nu este o prioritate.

Pe termen scurt, prioritățile sunt menținerea și actualizarea platformei și creșterea notorietății și a gradului de utilizare. Există planuri de a adăuga module și funcționalități noi, precum și o interfață mai bună pentru mobil, dar depind de identificarea unor finanțări.

II. DATE UTILIZATE ȘI PRODUSE

R4 folosește în principal date din surse publice, de pe data.gov.ro sau direct de la Ministerul Educației și instituțiile din subordine: învățământul profesional și tehnic, rețeaua școlară (inclusiv geocodări).

Pentru clasificări și căutări, R4 folosește nomenclatoare publice de la INS, Ministerul Educației și ONRC.

Pentru vizualizări de tip hartă, R4 folosește surse private non-profit, respectiv Open Street Map și geo-spatial.org.

Platforma nu generează date noi, ci îmbunătățește seturile de date folosite, prin normalizare, completare și curățare. Nu publică seturile de date îmbunătățite, în principal din rațiuni etice, dar sunt deschiși la a face acest lucru pe viitor.

III. RELAȚIA CU DATA.GOV.RO

La înființare, data.gov.ro a fost cel mai important furnizor de date pentru R4. Totuși, versiunile actualizate ale seturilor de date privind învățământul profesional și tehnic 2019-2020 au fost obținute direct de la furnizor, CNIDPT, și nu sunt încă disponibile pe portalul național de date deschise. Absența unui protocol cu furnizor ar fi putut duce la închiderea aplicației, pentru că data.gov.ro nu și-a respectat angajamentul de actualizare a datelor la timp.

Datele din data.gov.ro sunt descărcate în propriile servere și, în principiu, sunt îmbunătățite (curățare, geocodare, legare cu alte seturi de date). Datele îmbunătățite nu se întorc în portal.

Așteptări de la dezvoltarea portalului data.gov.ro:

- Publicarea unui calendar de actualizare a seturilor de date (măcar luna), care să facă predictibilă dezvoltarea aplicației;
- Respectarea angajamentelor privind actualizarea periodică a seturilor de date;
- Păstrarea structurii datelor la actualizare;
- Îmbunătățirea calității datelor: normalizare, curățare, formate moderne (ex.: JSON) sau măcar nu proprietare (ex. XLS);
- API-uri dedicate îmbunătățite;
- Folosirea consistentă a identificatorilor din nomenclatoarele oficiale în toate seturile de date;
- Noi seturi de date. Exemple:
 - Date INS (cel puțin nomenclatoarele);
 - Date ANAF, din declarația 112, spre exemplu numărul exact de angajați;
 - Date ANOFM: locuri de muncă disponibile la nivel național, actualizate frecvent, dacă nu în timp real;
- Un mecanism de feed-back îmbunătățit (vezi mai jos)

Echipa nu a folosit Disqus pentru feed-back, preferând comunicarea directă prin email. Actualul mecanism de feed-back nu este de încredere pentru că rareori conduce la un dialog cu furnizorul de date. Portalul data.gov.ro trebuie să medieze mai bine între furnizorii și utilizatorii de date. Să existe un mecanism care permite semnalarea erorilor punctuale din seturile de date, dar și obligația de a răspunde/acționa la semnalările primite.

IV. VALOAREA ADĂUGATĂ A PROIECTULUI

R4 are un istoric de aproximativ 14 luni. A avut aproape 8.000 de utilizatori unici și aproape jumătate de milion de vizualizări pe pagini. Timpul mediu de vizualizare este de aproape 3 minute/pagină, ceea ce sugerează că utilizatorii au identificat informații interesante. Proiectul are parteneriat cu instituții publice, dar și cu o platformă mass-media dedicată educației (edupedu.ro). S-au făcut cel puțin 2.000 de căutări folosind motorul intern.

Vârful de utilizare a fost în primăvara-vara 2020, atunci când, cel mai probabil, au fost luate decizii privind opțiunile elevilor. Se așteaptă creșterea utilizării în perioada similară a anului 2021.

NON-UTILIZATORI: U1, AUDITARE FINANCIARĂ

I. CE FACE U1

U1 este o mică firmă specializată în auditarea financiară a companiilor. Portofoliul include deopotrivă companii mijlocii și mari, U1 fiind prezentă pe piață de aproape 10 ani.

Utilizarea datelor deschise nu este partea principală a activității, dar reprezintă un complement absolut necesar în exercițiile de audit realizate. U1 auditează din punct de vedere al fluxurilor financiare clienții săi. Auditarea are însă nevoie de precizarea contextului în care evoluează clienții, cu alte cuvinte de diagnosticarea piețelor specifice pe care acești clienți activează.

Serviciile oferite de U1 se bazează astfel pe exploatarea intensivă și extensivă a datelor din domeniul public, fiind exploatate deopotrivă date furnizate de ANAF, INS, ANOFM, etc., precum și date readuse în domeniul public prin rafinare de către re-utilizatori precum R1-R4, sau rapoarte ale unor grupuri de consultanță majore (precum PWC sau E&Y) sau ale think-tank-urilor locale. Ansamblul acestor informații este rafinat în descrierea contextelor de piață ale firmelor auditate și în evaluarea performanței economice a competitorilor, ca element primordial în auditarea financiară a clienților și în dezvoltarea strategiilor de dezvoltare a acestora pe termen lung, mediu și scurt.

Produsul oferit de U1 este destinat exclusiv mediului de afaceri și este dedicat fiecărui client în parte, particularizările fiind continue.

II. DATE UTILIZATE ȘI PRODUSE

U1 folosește orice set de date din spațiul public care se potrivește obiectivului de activitate al clienților săi. Cu alte cuvinte, U1 este practic utilizator universal de date, și se află în căutare continuă a unor surse de date complementare.

U1 generează continuu date noi, dar acestea rămân în spațiul privat definit de relația cu clientul.

Nu sunt planuri de a trece la produse dedicate unei piețe generalizate sau a aducerii datelor în spațiul public (*shareability*), reutilizarea fiind singurul scop al activităților derulate.

III. RELAȚIA CU DATA.GOV.RO

Portalul data.gov.ro era complet necunoscut U1 la momentul realizării interviului, în decembrie 2020. Noutatea a fost completă, fiind însoțită chiar de surprinderea că un astfel de instrument există.

IV. VALOAREA ADĂUGATĂ A PROIECTULUI

Utilizarea datelor deschise este un element cheie în activitatea de auditare financiară realizată de U1. Fără îndoială, ea constituie doar o parte complementară în mixul auditării, dar importanța sa nu poate fi neglijată. Datele preluate din surse multiple sunt rafinate într-un produs complex și produc impact economic imediat. Unele dintre proiectele auditate sunt plasate în sfera economiei sociale, ceea ce face ca impactul concret al utilizării de date deschise prin activitatea U1 să se reflecte și în domeniul social.

UTILIZATORI ACADEMICI: CONCRETIZAREA ÎN PUBLICAȚII

Întrebarea la care răspunde acest demers, de studiere a publicațiilor ce utilizează data.gov.ro, este de a identifica dacă și cum este folosit data.gov.ro în mediul academic. Reutilizarea datelor constituie în acest sens un indicator de impact. Pentru a selecta articolele ce folosesc data.gov.ro am utilizat Google Scholar, un motor de căutare ce indexează articolele academice și cele din literatura gri (rapoarte de cercetare). Să notăm că aceasta este soluția cea mai puțin restrictivă: alte motoare de căutare (Scopus sau WoS, spre exemplu), iau în considerare doar aparițiile în jurnale de top, cu alte cuvinte cele de calitate bună și foarte bună. Utilizând Google Scholar am extins sfera, dat fiind și interesul pentru impact dincolo de mediul academic.

Căutarea³ a returnat 102 rezultate (cifra a fost actualizată pe 21.01.2021). Aceasta reprezintă deja un bun indicator de impact. Înseamnă că sunt 102 articole care folosesc într-un fel sau altul data.gov.ro ca element în construirea argumentului pe care îl propun.

Dacă acest impact este mic sau mare o putem afla prin a studia ce se petrece cu alte portaluri europene similare. Comparația cu alte țări trebuie însă realizată în funcție de mărimea acestora, dar și de alți indicatori, așa cum procedăm în capitolul dedicat analizei la nivel macro.

Pentru moment, să notăm doar că portalul românesc de date deschise, chiar și raportat la dimensiunea populației este mai performant decât cel din Bulgaria, Ungaria și Polonia și similar celor din Cehia sau Slovacia. Comparativ cu țări vestice, portalul italian are un impact mai redus decât cel românesc (prin ponderare cu dimensiunea populației), dar cele din Spania, Germania, Franța sau Olanda sunt mult mai folosite.

³ <https://scholar.google.com/scholar?q=data.gov.ro>

Tabelul 6. Impactul academic și în literatura gri a portalurilor de date deschise din EU, țările asociate și Moldova

Portal	Țara	Rezultate la căutarea pe Google Scholar	observații
data.gouv.fr	Franța	2010	
Data.gov.ro	România	102	
data.egov.bg	Bulgaria	5	
opendata.hu	Ungaria	9	
dane.gov.pl	Polonia	73	
dati.gov.it	Italia	362	
data.overheid.nl	Olanda	362	
datos.gob.es	Spania	636	
data.gov.hr	Croația	133	
data.gov.cz	Cehia	40	
data.gov.mt	Malta	8	
data.gov.sk	Slovacia	47	
podatki.gov.si	Slovenia	31	
data.gv.at	Austria	449	
govdata.de	Germania	579	
dataportal.se	Suedia	3	Suedia are două adrese pentru același portal
opnadata.se		51	
opendata.dk	Danemarca	44	
data.norge.no	Norvegia	190	
avoindata.fi	Finlanda	317	
vefur.island.is	Islanda	9	
opendata.swiss	Elveția	134	Elveția are două adrese pentru același portal
opendata.ch		132	
data.gov.be	Belgia	88	
dados.gov.pt	Portugalia	102	
data.gov.ie	Irlanda	235	
data.gov.uk	Marea Britanie	8390	
opendata.rtt.lt	Lituania	0	Lansat abia în 2020
avaandmed.eesti.ee	Estonia	0	Estonia are două adrese pentru același portal
opendata.riik.ee		40	
data.gov.lv	Letonia	23	
data.gov.gr	Grecia	137	
data.gov.cy	Cipru	27	
data.gov.md	Moldova	44	
data.gov.mk	Macedonia	1	
data.gov.rs	Serbia	39	
-	Turcia	-	Nu au portal
data.europa.eu	Uniunea Europeană	7390	

Vom reveni însă asupra acestor aspecte în capitolul dedicat analizei la nivel macrosocial.

Pentru moment, interesul este legat de utilizarea în sine a portalului românesc. Pentru a investiga calitativ în ce mod cele 102 articole produc un impact, am recurs la a selecta pe primele 11 raportate de Google Scholar și a le studia în detaliu. Acest lucru a făcut să luăm ca punct de referință pe cele mai accesate, dat fiind algoritmul de returnare a rezultatelor practic de motorul menționat.

Caseta 1. Lista articolelor academice studiate

1. Spoiala, M. N., Rinciog, O., & Posea, V. (2016, September). The semantic representation of open data regarding the Romanian companies. în *2016 15th RoEduNet Conference: Networking in Education and Research* (pp. 1-5). IEEE.
2. Rinciog, O., DOROBĂȚ, I. C., & Posea, V. (2017). ROUTE SUGGESTION FOR VISITING MUSEUMS USING SEMANTIC DATA. *eLearning & Software for Education*, 3.
3. Tănăsescu, L. G. (2017, July). Mechanisms underlying major investment projects în the Romanian public administration: strong and long term collaboration among stakeholders. în *Proceedings of the International Conference on Business Excellence* (Vol. 11, No. 1, pp. 949-959). Sciendo.
4. de Juana-Espinosa, S., & Luján-Mora, S. (2020). Open government data portals în the European Union: A dataset from 2015 to 2017. *Data în brief*, 29, 105156.
5. Badea, A. C., Badea, G., & Vasilca, D. (2019). THE ANALYSIS OF ROMANIA CADASTRAL RECORDS DYNAMICS în GIS USING OPEN DATA. *International Multidisciplinary Scientific GeoConference: SGEM*, 19(2.2), 713-722.
6. Lněnička, M., & Máchová, R. (2015, April). Open (big) data and the importance of data catalogs and portals for the public sector. în *Proceedings in Global Virtual Conference: The 3rd International Global Virtual Conference* (pp. 143-148).
7. Ghilic-Micu, B., Stoica, M., & Uscatu, C. (2015). Challenges of 4D (ata) Model for Electronic Government. *Informatica Economica*.
8. Rinciog, O., & Posea, V. (2015, November). Publishing Romanian public health data as Linked Open Data. în *2015 E-Health and Bioengineering Conference (EHB)* (pp. 1-4). IEEE.
9. Rinciog, O., & Posea, V. (2018). STATE OF ROMANIAN OPEN DATA în 2017. *UNIVERSITY POLITEHNICA OF BUCHAREST SCIENTIFIC BULLETIN SERIES C-ELECTRICAL ENGINEERING AND COMPUTER SCIENCE*, 80(4), 3-12.
10. Rinciog, O., & Posea, V. (2016). LODRo: using cultural Romanian open data to build new learning applications. în *The International Scientific Conference eLearning and Software for Education* (Vol. 1, p. 267). "Carol I" National Defence University.
11. Burloiu, V. P. (2016). Publishing Shame Lists on the Internet: Respecting Data Protection Principles vs Public Interest. *Eur. Data Prot. L. Rev.*, 2, 403.

Pentru aceste 11 texte, listate în Caseta 1, am urmărit câteva criterii relevante prin prisma obiectului studiat. În primul rând am identificat obiectivul studiului raportat în textul respectiv. Am constatat astfel că reutilizarea datelor spre oferire în alte aplicații este prezentă în trei cazuri, folosirea portalului ca sursă pentru analize academice este prezentă într-un singur produs, iar meta-analiza este elementul comun a șapte din cele 11 texte, în care portalul este de fapt obiect de studiu în sine și nu o sursă de date.

Al doilea element relevant este producția de noi date care să poată fi la rândul-le reutilizate. Acest lucru este realizat în 3 din cele 11 texte studiate.

Al treilea element de interes este impactul acestor articole. Ar fi fost util să avem acces la altmetrics: câte vizualizări și descărcări a avut fiecare articol. Astfel de informații nu sunt însă disponibile sau atunci când sunt disponibile nu sunt măsurate într-un mod similar. În absența unor altmetrics comparabile, am recurs la indicatorul standard de observare a impactului unui produs academic și anume numărul de citări. Acesta este mai degrabă redus în majoritatea cazurilor, dar este în acest caz și efectul vârstei tinere a lucrărilor luate în considerare. Timpul poate schimba acest aspect, probând calitatea celor 11 materialele.

În al patrulea rând, am căutat să aflăm și în ce măsură textele care utilizează data.gov.ro sunt parte din *open science*, investigând un singur aspect în acest sens: dacă există acces deschis la produsul respectiv. Cu trei excepții, accesul la publicațiile respective este gratuit.

Tabelul 7. Tabel sinoptic al analizei asupra articolelor luate în considerare

paper	scop	Produce date noi?	Număr citări	Acces open?	Jurnal/Editură de top?
1	Portalul este obiect de studiu	nu	2 (internaționale)	Nu	Nu
2	Reutilizare a datelor despre muzee și oferire de servicii publicului	da	3 (românești)	Da	Nu
3	Analiză de date data.gov.ro	nu	1 (internațional)	Da	Nu
4	Descriere a portalurilor de date deschise	nu	2 (internaționale)	Da	Da
5	Analiză și reutilizare de date din data.gov.ro	da	1 (internațional)	Da	Nu
6	Descriere a portalurilor de date deschise	nu	11 (internaționale)	Da	Nu
7	Portalul este obiect de studiu	nu	6 (internaționale)	Da	Nu
8	Portalul este obiect de studiu	nu	1 (autocitare)	Da	Nu
9	Portalul este obiect de studiu	nu	0	Da	Nu
10	Reutilizare - muzee	Da	5 (internaționale)	Nu	Nu
11	Portalul este obiect de studiu	nu	0	Nu	Da

În fine, calitatea textului respectiv este dată și de calitatea locului unde a apărut. Doar câteva dintre aceste produse sunt publicate în jurnale sau la edituri cu peer-reviewing și doar două sunt susținute prin a fi parte a unor jurnale de top. Majoritatea textelor sunt parte a unor publicații românești de flux secund în știință sau sunt materiale prezentate în conferințe. Ele pot fi desigur de calitate, dar pentru moment timpul scurt nu a permis probarea acestui fapt prin citări.

Evaluarea rezultatelor de mai sus poate fi realizată în contextul internațional al literaturi despre date deschise. Google Scholar raportează 352.000 de rezultate la căutarea după „open data”. Căutări similare în jurnale de top pot fi realizate folosind Scopus (11.695 rezultate) sau Web of Science (6.884 rezultate). Numărul ridicat de rezultate arată că subiectul este de interes în lumea academică și constituie un argument pentru demersul de față. Poziționarea bună a data.gov.ro în raport cu alte portaluri europene, în condițiile în care România investește cel mai puțin în cercetare, este un indiciu privind succesul portalului românesc de date.

Calitatea articolelor ce utilizează data.gov.ro ca sursă poate fi judecată prin raportare la meta-studii similare acestei secțiuni. Corrales-Garay et al (2019) au trecut în revistă articolele din WoS și Scopus care pun în relație „open data” și „open innovation”. Diferența majoră față de cele 11 articole analizate de noi este că două treimi dintre cele studiate de Corrales-Garay et al (2018) erau de natură empirică (adică exploatau date, nu doar discutau despre deschiderea datelor), fiind centrate pe reutilizare de date (deschise). Din acest punct de vedere, impactul data.gov.ro este încă modest la nivel internațional, dar comparat cu prezența încă și mai modestă a cercetării românești în cea internațională, evaluarea devine în fapt pozitivă.

Sintetizând, avem, de a face cu un flux mediu la nivel european de articole ce folosesc data.gov.ro, dar peste cel observat la țările similare ca dezvoltare României. Din păcate, nu există indicii certe privind calitatea acestui impact și a impactului produselor rezultate, iar așteptările în acest sens sunt mai degrabă modeste. Există însă premisele ca să asistăm la un proces prin care cantitatea să facă loc în timp și calității.

CONCLUZII

TIPURI DE IMPACT

Studiile de caz vorbesc în principal despre impactul potențial al reutilizării datelor furnizate de portalul național data.gov.ro, dar și despre limitările variantei curente a portalului.

Se poate observa direct impactul economic (R3, afacere sustenabilă după un an de la lansare), social (proiecte non-profit cu un *reach* bun în publicul țintă) sau chiar politic și de politici (R2, proiect asumat de un partid politic parlamentar, care a inclus elemente din el în propriile propuneri de politici publice). Chiar dacă acest tip de impact nu poate fi cuantificat imediat, el este cert.

Utilizarea academică este redusă și este plasată mai degrabă în jurnale cu factor de impact mic. Nu am identificat studii de caz care să țină seama de impactul de mediu.

În niciunul din studiile de caz analizate nu putem vorbi de impact în cascadă. Deși utilizatorii generează noi seturi de date, acestea nu sunt deschise.

PUNCTE TARI: DISPONIBILITATE, NOTORIETATE ȘI LEGITIMITATE

Data.gov.ro are ca impact evident o aducere în spațiul public a datelor, ceea ce a fost apreciat de toți cei care utilizează portalul preluând informație din el. Este posibil să asistăm în fapt și la un efect de multiplicare, prin preluarea exemplului de alte organizații ale administrației publice. Aceasta probează în sens pozitiv legitimitatea data.gov.ro.

Unul dintre respondenți a punctat: „noi avem un om care, periodic, verifică ce date noi au mai fost publicate, în căutarea unor idei noi de afaceri”. În opinia tuturor reutilizatorilor, portalul de date deschise se bucură de legitimitatea necesară pentru a atrage atenția dezvoltatorilor.

Notorietatea portalului este considerată în general în creștere pentru publicul larg, dar bună și foarte bună în comunitatea specialiștilor în manipularea datelor. Acest lucru nu se confirmă însă în ce privește comunitatea academică.

PUNCTE SLABE: ACTUALIZARE, COMPLETITUDINE, EXTINDERE

Calitatea datelor este chestionată de toți re-utilizatorii ce aduc date în spațiul public sau prin afacerile pe care le dezvoltă

În aproape toate studiile de caz analizate, utilizatorii s-au văzut nevoiți să completeze datele din portalul național de date deschise cu alte surse, inclusiv pentru seturi de date care sunt deja prezente pe portal. Astfel de exemple sunt datele privind achizițiile publice sau cele din educație. În ambele situații, utilizatorii preferă să folosească scripturi sau alte metode neconvenționale pentru a descărca datele direct de la furnizori, chiar dacă nu sunt publicate structurat, pentru că pe portalul de date deschise nu sunt actualizate cu aceeași frecvență. În alte situații, cum sunt datele privind execuția bugetară sau datele INS, seturile respective nici măcar nu ajung pe data.gov.ro. În fine, o a treia situație, cel mai bine exemplificată de ONRC, setul de date publicat este limitat și nu include câmpuri importante. Această „migrare” de la portalul de date deschise către furnizorii originali, instituții publice, subliniază trei factori care limitează impactul portalului:

- Actualizare: unele seturi de date nu sunt actualizate cu o frecvență și predictibilitate acceptabile;
- Completitudine: seturi de date esențiale nu sunt complete;
- Extindere: instituții publice care dețin seturi de date importante nu publică pe portalul de date deschise.

LIMITĂRI: CALITATE ȘI PRELUCRARE AUTOMATĂ, API, FEED-BACK

În toate cazurile de reutilizare, datele din portal sunt duplicate pe serverele utilizatorilor și nu utilizate așa cum sunt oferite de data.gov.ro. Sunt mai multe tipuri de limitări care conduc la această situație:

- Toate seturile de date reutilizate – practic, toate seturile de date cu valoare ridicată – au o calitate scăzută. Este necesară curățarea datelor;
- Numeroase seturi de date sunt în formate precum XLS, care oferă un potențial redus de prelucrare automată;
- API-ul pus la dispoziție de portal este neprietenos și este privit cu neîncredere.

O posibilă reîntoarcere a datelor noi, sau măcar îmbunătățite, în domeniul public ar putea fi feed-back-ul utilizatorilor. Spre exemplu, corecturi punctuale ale unor seturi de date pot fi integrate în setul original de date. Niciunul dintre utilizatori nu s-a arătat însă interesat de instrumentul pus la dispoziție – forumul de tip Disqus. Calitatea scăzută a interacțiunii cu utilizatorii reprezintă o limitare a impactului potențial.

AȘTEPTĂRI: NOI SETURI DE DATE

Fiecare dintre observațiile privind punctele slabe și limitările implică așteptări privind calitatea interacțiunii cu data.gov.ro, respectiv:

- Creșterea calității datelor, în sensul arătat mai sus;
- Îmbunătățirea API;
- Mecanisme de feed-back mai bune.

Așteptarea cea mai importantă este însă și cea evidentă: noi seturi de date, în principal seturi cu valoare adăugată mare. Numeroase exemple au fost date de respondenți și sunt incluse în prezentările studiilor de caz.

Nu în ultimul rând, portalul trebuie să își construiască o punte de legătură cu comunitatea academică, pentru a putea vorbi de impact în cercetare și educație.

PANELUL DELPHI

Rolul panelului Delphi a fost de a evalua impactul data.gov.ro pe baza estimărilor produse de experții din domeniu. Am argumentat în introducerea acestui volum despre utilitatea demersului, care constituie de altfel un element de bază și în alte evaluări ale deschiderii datelor. Rolul experților este esențial: ei sunt cei care au, cel puțin teoretic, cea mai bună cunoaștere asupra subiectului și care pot evalua prin prisma propriei experiențe directe de lucru stagiul deschiderii datelor în România. Fără îndoială, perspectiva este una incompletă, motiv pentru care analiza rezultatelor panelului Delphi este doar o parte a informației pe care acest volum o folosește în demersul de evaluare a deschiderii datelor din România.

METODA

Metoda presupune propunerea unor teme de discuție și solicitarea unui grup de experți să se pronunțe în mod iterativ asupra acestora. Mai exact, propunerea temelor inițiale este urmată de comentarii asupra acestora, apoi de comentarii la comentarii și așa mai departe.

În cazul de față am optat pentru un Real-Time Delphi, pe care l-am susținut logistic prin intermediul platformei Kialo. În colaborare cu SGG, autorii acestui volum au selectat experți activi în domeniul datelor deschise prin participare la acțiuni legate de acestea, prin implicare efectivă în deschiderea datelor, utilizarea și reutilizarea acestora și prin poziționare în organizații cheie ale administrației publice cu rol în deschiderea datelor.

Funcționarea consultării este relativ simplă: au fost propuse un set de teme sau enunțuri cheie, iar participanții au fost încurajați să comenteze asupra lor, adăugând argumente în favoarea acestora (*pros*) sau care să le pună sub semnul întrebării (*cons*).

Fiecare participant putea adăuga cons și pros la comentariile oricărui altul, iar răspunsurile se pot acumula într-un arbore al discuției practic nesfârșit. La final, am prevăzut o rundă în care participanții puteau vota oricare dintre comentarii.

Au fost propuse patru teme de discuție majore, urmând structura cercetărilor ce stau la baza acestui volum: notorietatea, legitimitatea, oferirea de date, și utilizarea datelor. Ele au fost propuse participanților prin textul din Caseta 2. Textul a oferit pe scurt informații privind obiectivul exercițiului, a detaliat conținutul celor patru teme și a explicat ce are fiecare de făcut și în ce constrângeri de timp. Portalul a fost însă lăsat deschis o vreme mai îndelungată, tocmai pentru a permite revenirea cu comentarii.

Invitațiile de a participa la discuție au fost trimise în decembrie 2020, discuția având loc pe o perioadă de trei săptămâni, înainte de Crăciun. Un ultim comentariu a fost adăugat de participanți la finalul lui ianuarie 2021.

Pentru a simplifica participarea și a stimula activitatea, au fost propuse de la bun început câteva comentarii inserate de către autorii acestui studiu:

- La „oferirea de date”, în secțiunea de *pros* a fost adăugat enunțul „data.gov.ro a crescut cu mult oferta de date deschise din RO prin faptul că a furnizat un exemplu potențialilor furnizori”.
- La „utilizare”, am inserat un *con*: „data.gov.ro nu oferă acces la date printr-un API, ceea ce face dificilă reutilizarea seturilor de date publicate”

- La „notorietate” am propus un *con* extrem, menit să stârnească reacții: “data.gov.ro nu are niciun impact, nimeni nu știe de portal!”

Caseta 2. Prezentarea discuției așa cum a fost ea afișată pe Kialo experților invitați ca participanți

Data.gov.ro a devenit funcțional în 2013.

Între timp am trecut printr-un proces de dezvoltare, atât la nivel de acumulare de date, cât și de mod de diseminare, utilizare, calitate a datelor etc.

Ceea ce dorim în acest exercițiu este să evaluăm impactul data.gov.ro (dar și al întregului proces de deschidere a datelor) prin apelul la opiniile experților români din domeniu.

Vă invităm să comentați și să reacționați la comentariile celorlalți în 2-3 runde succesive pe temele de mai jos.

Ideea este simplă.

Noi lansăm 4 teme mari în care se poate observa sau nu un impact:

- NOTORIETATEA data.gov.ro, ca prim element ce poate conduce la impact
- LEGITIMITATEA (măsura în care deschiderea datelor este considerată firească de producători și utilizatori potențiali)
- OFERIREA DE DATE deschise, atât prin data.gov.ro, cât și efectele de multiplicare datorate preluări modelului de producători ce nu distribuie (și) prin data.gov.ro
- UTILIZAREA efectivă (incluzând reutilizarea) și consecințele acesteia

Fiecare din cele 4 teme pot fi privite ca vectori de impact. Impactul se poate manifesta în domeniul (1) social, (2) economic, (mediu), (4) politic și administrație. (clasificarea o regăsim ca atare în mai multe lucrări, incluzând Open Data Maturity Report 2019 (Blank, 2019)).

Prima rundă de discuții are ca scop generarea unui set de afirmații privind impactul. Vă rugăm să vă exprimați așa cum doriți privind impactul data.gov.ro pe cele 4 dimensiuni și să comentați la ce spun ceilalți.

La finalul rundei, vom extrage afirmațiile care acoperă un spectru larg de impact și supunem comentariilor doar acest set de afirmații, așa cum reiese el din discuție. Dacă simțim că deja discuțiile au fost suficiente, sărim direct la runda a treia în care se votează afirmațiile.

Runda a treia poate elimina unele afirmații prin vot negativ și le ordonează. Avem, astfel o evaluare expert a ceea ce s-a petrecut în România cu deschiderea datelor (avem și alte 5 metode alternative de evaluare a impactului. În absența baseline study, evaluarea devine validă prin triangulare).

PS. se poate reacționa la comentariile celorlalți și prin a vota ceea ce propun ei ca afirmații, și comentând în thread-urile respective. (votul va fi însă permis doar în ultima rundă).

Mod de folosire

=====

Click pe una dintre cele 4 direcții propuse.

Puteți adăuga un comentariu fie la pros, fie la cons.

Apoi puteți adăuga comentarii la comentarii.

Șamd

termene:

=====

Prima rundă de comentarii se încheie joi, 17 decembrie.

A doua e scurtă: până luni, 21 decembrie (ora 12).

Apoi lăsăm 2 zile pentru a vota afirmațiile ce au apărut ca rezultat al discuției.

RATE DE RĂSPUNS

Au fost selectați 36 de potențiali experți. Aceștia au fost invitați prin email să ia parte la discuția de pe Kialo. 18 dintre aceștia nu au accesat link-ul, nerăspunzând astfel invitației. 18 au accesat discuția, generând 605 vizualizări, 9 au participat efectiv la adăugarea de comentarii. 32 de comentarii distincte au fost adăugate, între două și șase pentru fiecare expert în parte.

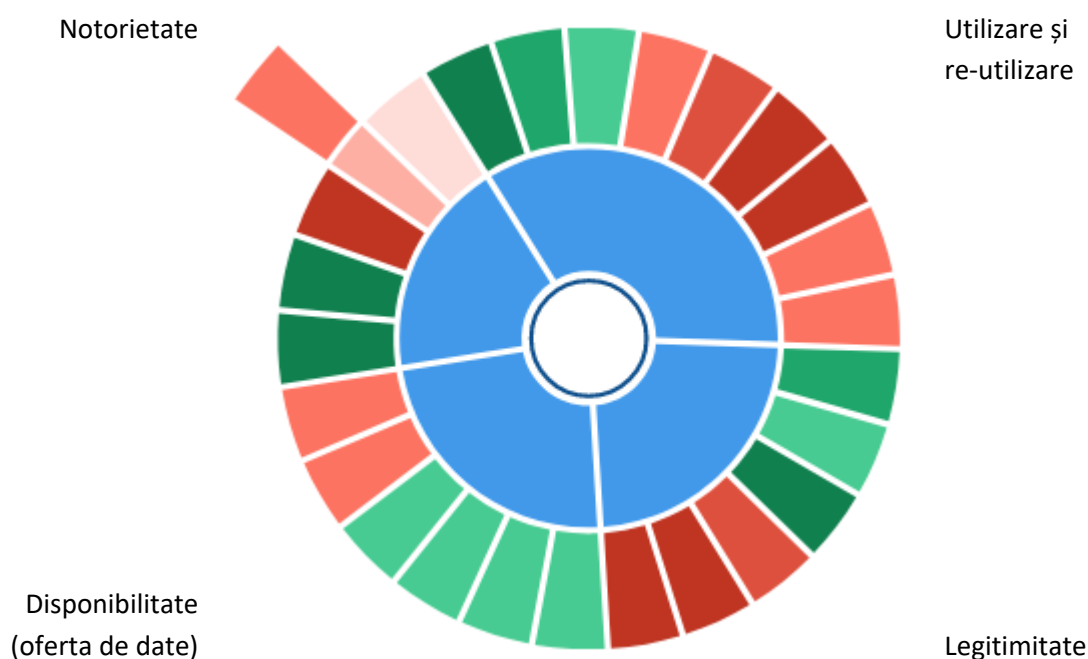
Ratele de răspuns nu sunt mari, dar indică un interes moderat față de temă. Ținând cont și de contextul nepotrivit al derulării exercițiului (pandemia, finalul de an cu finalizarea de proiecte pentru mulți dintre cei implicați, și sărbătorile de iarnă), avem de fapt de a face cu o rată de răspuns destul de bună.

Calitatea comentariilor, pe care le discutăm în secțiunea următoare, este și ea ridicată, constituind un alt indicator pentru relevanța concluziilor pe care le tragem pe baza acestui exercițiu.

TIPURI DE REACȚII

Figura 5 ilustrează modul în care a decurs dezbaterea. Cele patru teme majore (petele albastre) au atras și reacții de respingere (marcate cu colori apropiate de roșu), denotând o chestionare a impactului produs, și reacții de aprobare (petele verzi). Numărul reacțiilor verzi și a celor roșii este echilibrat, denotând o abordare echilibrată asupra impactului, așa cum a rezultat el din interacțiunea grupului de experți. Există un singur argument adus unuia dintre argumentele de nivel doi. Acel argument este contra unui argument contra, deci practic, prin semnificație, este „verde” nu „roșu”. În afara argumentelor prezentate în topologie, mai sunt și comentarii la argumente, pe care experții participanți le-au inserat însă nu ca opțiuni contra sau pro, ci ca simple comentarii. Le detaliem pe toate în secțiunea următoare.

Figura 5. Topologia discuției Delphi: cu albastru sunt temele inițiale, cu verde sunt argumente pro, iar cu roșu argumente contra



REZULTATELE

NOTORIETATEA

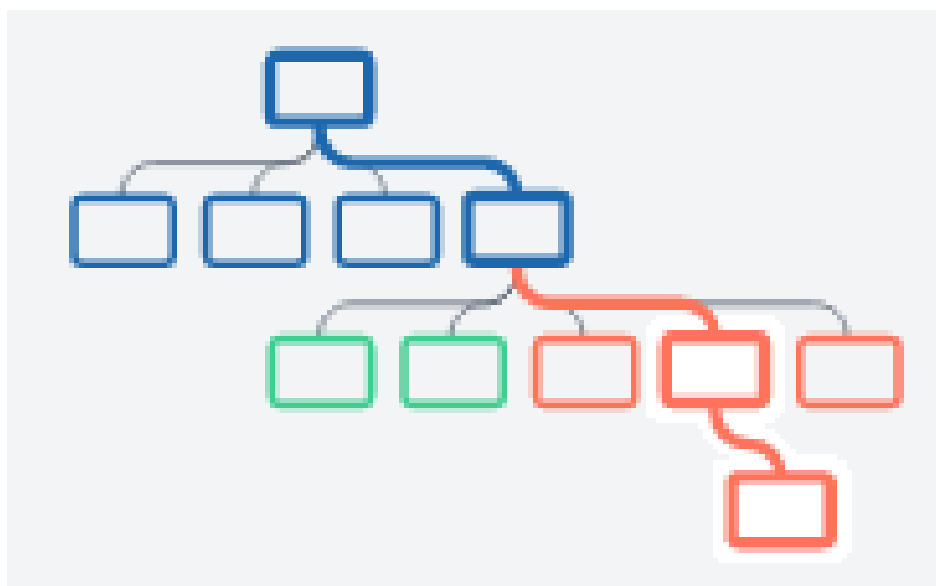
Pentru a înțelege mai bine modul de adăugare a argumentelor, Figura 6 ilustrează procesul înfățișând structura sa ierarhică. În partea de sus, pătratul albastru singur este tema majoră a exercițiului „impactul data.gov.ro”. Cele patru direcții subsumate – notorietate, legitimitate, disponibilitate, utilizare – constituie nivelul imediat următor. Apoi, sub fiecare direcție sunt argumentele pro și contra. Spre exemplul, figura indică trei argumente contra sub notorietate (roșii) și două pro (verzi). Unul dintre argumentele contra are la rândul său un contraargument.

Tabelul 8. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Notorietatea

Impactul în ce privește NOTORIETATEA (există? ce ați observat? ce se mai putea face?)	
Comentarii primite: Mulți oameni de IT de pe lângă instituții, atât contractori cât și angajați, spun că un scop final este publicarea datelor pe acest portal. Chiar dacă nu toți înțeleg ce este implicat în asta, am văzut de mai multe ori că siteul e dat ca referință. Deci, din punct de vedere notorietate, nu văd multe probleme. Pe de altă parte, din datele mele, dezvoltatorii potențiali consumatori nu iau în calcul serios să folosească datele din site pentru construirea de aplicații din lipsa unui set consistent de date. Dar notorietate pare să existe în mediile IT.	
PROS	CONS
In mediile în care s-a reușit promovarea și implicarea este cunoscut. Notorietatea sa ca portal trebuie completată de cel puțin două componente - povești de succes \oamenii din administrație și din afară au nevoie de exemple concrete\ și leadership \ideal la un nivel destul de înalt\.	<p><u>PREDEFINIT:</u> data.gov.ro nu are niciun impact, nimeni nu știe de portal!</p> <p><u>Comentarii primite:</u></p> <ol style="list-style-type: none"> 1. nu ai dreptate "data.gov.ro" are un impact dar mult mai mic decat ce ar trebui/este în alte tari sa fie și nu se poate spune ca nu are nici un impact. 2. Are impact doar între cunoscătorii domeniului datelor deschise. 3. atât timp cât datele de pe data.gov.ro ajung pe portalul european nu este deloc relevant câte persoane accesează portalul românesc. Poate ca multi prefera interfata celui european dar asta nu înseamnă ca cel european ar putea exista fara cel românesc (sau oricare alt portal national). Situatia este similara cu cea a Eurostat / INS. <p>Nu sunt cunoscute prea multe exemple \eufemistic vorbind\ ale unor aplicații sau servicii care să folosească seturile de date de pe data.gov.ro. Asta se leagă și de lisa generală de vizibilitate a portalului.</p> <p>Modelul de business bazata pe consumul datelor deschise implica ascunderea sursei datelor / nementiunarea portalului. Astfel se îngreunează apariția unei oferte concurente chiar dacă este încălcată obligația de atribuire din licență</p> <ul style="list-style-type: none"> • CON: Open data și implicit data.gov.ro nu este perceput ca un instrument major al guvernului în întregul sau - nu este referit pe site-ul de web al guvernului \www.gov.ro și nici măcar la utile sau media\ și nici pe site-urile ministerelor. De exemplu vezi site-ul (https://www.gouvernement.fr/) unde jos este link-ul rapid către http://data.gouv.fr/
Portalul este menționat de legi și norme	

Tabelul 6 prezintă opiniile experților consultați privind notorietatea data.gov.ro, redând toate argumentele și reacțiile pro și contra la ele, precum și comentariile formulate. Imaginea de ansamblu este una ușor pozitivă.

Figura 6. Structura arborescentă a răspunsurilor, cu detaliere pe ramura „notorietății”



*"Ramura notorietății" este cea din dreapta, cu căsuțele îngroșate. Practic, căsuța de sus de tot reprezintă tema mare, rândul căsuțelor cu albastru este diferențierea pe cele patru teme mari, dintre care notorietatea este cea mai din dreapta. „Notorietatea” a primit două comentarii pozitive (*pro*) – și anume cele cu verde, și trei *contra* (negative) – cele cu portocaliu. Unul dintre comentariile negative a primit la rândul său o reacție *contra*.

Argumentul negativ predefinit (*data.gov.ro nu are niciun impact, nimeni nu știe de portal!*) a stârnit reacții de respingere parțială. Experții par să conchidă în cele trei comentarii că există o anumită notorietate, ce e drept limitată la un cerc de inițiați și sub potențial. Al treilea comentariu, mai amplu, indică faptul că în fapt utilizarea poate fi realizată prin alte instrumente, ceea ce face ca notorietatea *data.gov.ro* să fie doar un element mai puțin relevant în evaluare.

Al doilea argument negativ, în fapt primul formulat de experții consultați (*Nu sunt cunoscute prea multe exemple \eufemistic vorbind\ ale unor aplicații sau servicii care să folosească seturile de date de pe data.gov.ro. Asta se leagă și de lisa generală de vizibilitate a portalului.*), face trimitere la notorietate prin utilizare și aduce în sine o nevoie de monitorizare a utilizării și de popularizare a site-ului.

Al treilea *con* din serie, cel care la rândul său a stârnit un alt *con*, aduce în discuție utilizarea neraportată a portalului guvernamental, ceea ce de fapt ar îngreuna observarea corectă a notorietății. Răspunsul la această obiecție (care în fapt este un argument în favoarea notorietății) conduce la o nevoie de abordare integrată și la absența sprijinului de imagine complet al guvernului. Este adus în prim-plan exemplul francez, reliefându-se absența implicării aparatului guvernamental în promovarea *data.gov.ro* prin site-urile sale.

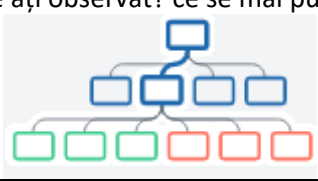
Primul argument *pro* merge în aceeași direcție, extinzând în fapt nevoia de implicare guvernamentală până la nivel de leadership.

Al doilea argument *pro* (*Portalul este menționat de legi și norme*) asociază impactul în sfera notorietății cu un instrument subtil de promovare și recunoaștere: prezența în reglementările curente ce definesc funcționarea statului român. Aceasta aduce *data.gov.ro* în atenția constantă a unor audiențe diverse și îl promovează permanent.

LEGITIMITATEA

Tabelul 9 listează enunțurile oferite de participanții la discuție cu privire la legitimitate. Tema dominantă este legătura cu alte portaluri (ale administrației publice sau ale sistemului public în general) și cooperarea cu acestea, în condițiile în care data.gov.ro constituie elementul central al acestei arhitecturi complicate.

Tabelul 9. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Notorietatea

Impactul în ce privește LEGITIMITATEA (există? ce ați observat? ce se mai putea face?)	
	
PROS	CONS
<p>data.gov.ro beneficiază de legitimitatea dată de domeniul gov.ro - nu e de neglijat acest aspect. Gov.ro aduce și partea de „oficial” dar și componentele de seriozitate atribuite unei instituții publice.</p>	<p>data.gov.ro nu e singurul portal care agregă seturi de date de la instituții publice. Transparența bugetara, de pildă, conține o mulțime de date despre execuțiile bugetare ale instituțiilor publice, dar cele două platforme par să nu aibă nicio legătură una cu alta.</p>
<p>Clar are legitimitate dar din pacate nu cred ca și intre institutii în sine. Daca au nevoie de documente, cred ca în continuare isi cer între ele direct, nu se uita pe portal daca nu cumva sunt publicate deja.</p> <p><u>Comentarii primite:</u></p> <ol style="list-style-type: none"> 1. Cred că această situație este de fapt un contrargument privind legitimitatea portalului. 2. Ce înseamnă legitimitate ? în sens strict - adică se conformează legilor în vigoare - are deplina legitimitate. Dacă sensul este de "sursa oficială de date", cum vada că și-ar dori [autorul argumentului de mai sus], atunci aceasta misiune este în directă contradicție cu principiul Open Data. Datele deschise sunt libere de orice constrângere, inclusiv constrângeri ale celui care le publica. Este chiar menționată explicit în legislație ideea ca publisherul nu poate fi obligat să țină datele la zi. Ori dacă datele nu sunt ținute la zi în mod evident acestea nu sunt corecte iar dacă nu sunt corecte atunci nu pot fi folosite în proceduri administrative sau orice altceva înafara unor prelucrări statistice sau fără consecințe directe. 	<p>se pare ca nu exista vointa la nivel local sa se utilizeze data.gov.ro preferand construirea unor platforme proprii.</p> <p><u>Comentarii primite:</u></p> <ol style="list-style-type: none"> 1. Consider că impactul pe care îl are proliferarea platformelor locale pentru publicarea datelor deschise - mai ales în lipsa unei standardizări a modului în care funcționează (cum se publică datele, de către cine, modul în care utilizatorul interacționează cu platforma,) - este major în lipsa unui mod automatizat de a prelua datele publicate astfel pe data.gov.ro. 2. Din câte știu, portalurile locale au gândit și o metodă automată de schimb de date cu data.gov.ro. Nici nu ar fi logic altfel.
<p>Pentru mine, ca antreprenor ce dezvoltă o soluție bazată pe date publice, data.gov este primul loc în care caut setul de date de care am nevoie. Este mult mai ușor de urmărit un singur portal centralizator, decât mai multe și mici. Deci trebuie insistat în a aduce cât mai mulți furnizori de date în portal.</p>	<p>data.gov.ro trebuie să folosească legitimitatea dată de faptul că este portal național pentru a integra și alte portaluri locale. Nu e o competiție cu acele portaluri - data.gov.ro poate chiar promova aceste exemple cu condiția integrării datelor și în data.gov.ro</p>

Legitimitatea este indiscutabilă în opinia experților, dar are câteva puncte critice ce o pot pune în pericol dacă nu sunt abordate atent. Legitimitatea derivă din poziționarea în spațiul guvernamental și din viziunea integratoare. (primul și al treilea comentariu *pro*). Actualizarea permanentă a datelor constituie o condiție sine qua non. De aici vin și riscurile. Proliferarea surselor de date din afara data.gov.ro este văzută ca un pericol la adresa platformei guvernamentale, care pierde din misiunea integratoare.

Este interesant de observat că implicația pare a fi cerința ca data.gov.ro să includă toate datele furnizate de alte entități. În fapt nu este așa, ci ar putea fi suficient un motor care să ofere link-uri către restul furnizorilor, cu condiția unei indexări care să includă și toate metadata necesare.

DISPONIBILITATEA (OFERTA DE DATE DESCHISE)

În ciuda comentariilor plasate mai degrabă în zona de *pros*, după cum relevă Tabelul 9, lectura acestora sugerează rezerva experților în ce privește oferta de date de pe data.gov.ro. Aceasta este privită mai ales prin prisma utilizării (care în fapt constituie obiectul secțiunii următoare, nu a acesteia). Argumentele prezentate trimit mai degrabă la nevoia unui studiu de impact precum acesta și a unei monitorizări permanente însoțită de raportări periodice în ce privește gradul de utilizare, numărul de seturi noi oferite, a celor actualizate, a exemplelor de utilizare etc.


Tabelul 10. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Notorietatea

Impactul în ce privește OFERIREA DE DATE (există? ce ați observat? ce se mai putea face?)	
PROS	CONS
<p><u>PREDEFINIT:</u></p> <p>data.gov.ro a crescut cu mult oferta de date deschise din RO prin faptul că a furnizat un exemplu potențialilor furnizori</p>	<p>Numărul seturilor de date oferite a crescut foarte lent în ultimii ani.</p>
<p>data.gov.ro are un potențial mare de creștere dacă este ajutată și de o politică în domeniu. Obligarea/urmărirea autorităților de a publica ce fel de date gestionează ar fi un bun start</p>	
<p>Pentru că alegerile locale s-au încheiat de curând, este un bun moment să fie ținute cursuri pentru primării. Datele lor ar fi o adiție valoroasă la portal.</p>	
<p>Îmbunătățirea portalului - o interfață mai prietenoasă, secțiuni noi etc. - reprezintă un avantaj important pentru modul în care datele publice deschise sunt oferite către public.</p>	<p>Putem judeca oferirea de date după folosirea lor. Nu știu aplicații care să se bazeze pe datele de pe acest portal. O analiză ar mai putea fi făcută și după periodicitatea cu care se pun datele la dispoziție - dacă se respectă ciclicitatea declarată, e foarte bine, pentru că antreprenorii pot avea predictibilitate. Un alt mod de măsurare poate fi după diversitatea tipurilor de instituții care expun date deschise. Din punctul meu de vedere, date de la instituțiile locale sunt mult prea puține.</p>

UTILIZAREA ȘI RE-UTILIZAREA

Discuția despre (re)utilizare este dominată de opiniile negative, atât numeric cât și ca substanță. Opiniile pozitive sunt mai degrabă generale, despre creșterea numărului de date din portal, nu neapărat despre modul cum sunt utilizate. A fost dat un singur exemplu de reutilizare, punctual, într-un context special. În schimb, opiniile negative indică clar punctele slabe ale portalului: nu sunt publicat suficient de multe seturi de date cu valoare economică ridicată, iar datele existente au o calitate scăzută, dată de lipsa standardizării, frecvența scăzută a actualizării (sau absența unui angajament privind actualizarea) și absența metadatelor. Argumentele prezentate sugerează că reutilizarea scăzută este determinată (și) de aceste probleme structurale ale portalului, și că acestea ar trebui abordate prioritar.

Tabelul 11. Reacțiile panelului de experți asupra impactului data.gov.ro în ce privește Utilizarea și re-utilizarea

Impactul în ce privește UTILIZAREA și RE-UTILIZAREA (există? ce ați observat? ce se mai putea face?)	
	
PROS	CONS
data.gov.ro a crescut cu mult exemplele de utilizare a datelor deschise	<u>PREDEFINIT:</u> data.gov.ro nu oferă acces la date printr-un API, ceea ce face dificilă reutilizarea seturilor de date publicate
	Calitatea seturilor de date este foarte inegală \ (asta ține și de lipsa standardizării în ceea ce privește colectarea datelor în sectorul public\).
numarul de seturi de date a crescut	nu toate seturile existente sunt actualizate, facand utilizarea acestora irelevantă în contextul dinamicii actuale
	Până la aspectele tehnice, trebuie construită o platformă de încredere că datele existente sau viitoare vor fi deschise permanent, vor crește în calitate - acel open by default ca angajament al administrației.
Datele deschise privind achizițiile publice au fost una dintre sursele principale ale anchetei jurnalistice privind Hexipharma. Este unul dintre cele mai de impact exemple ale utilizării datelor deschise.	Nu avem seturi de date cu valoare economică mare. Exemplu: date de trafic, date legate de firme etc
	În multe situații lipsește descrierea setului de date și descrierea coloanelor, lipsește acel fișier de metadata care să-mi spună într-un mod foarte descriptiv la ce mai uit.

CONCLUZII ȘI IMPLICAȚII

Cele patru teme prezintă numeroase întrepătrunderi și interdependențe. Analiza lor separată a permis identificarea specificităților, dar conduce și la ideea integrării în acțiunile de urmat a unor elemente legate de fiecare dintre teme.

Notorietatea este una limitată, dar aceasta poate fi datorată pe de o parte lipsei de promovare, iar pe de alta poate fi o falsă impresie dată de monitorizarea dificilă a utilizării portalului guvernamental. O primă recomandare vine derivă din comentariile experților consultați: Guvernul României poate promova data.gov.ro prin site-urile pe care le gestionează, atât la nivel ultra-central (gov.ro), cât și prin site-urile ministerelor și agențiilor administrației publice centrale. În prezent, acest lucru lipsește din mixul promovării politicii deschiderii datelor și poate afecta inclusiv legitimitatea demersului.

Legitimitatea este implicită pentru cei aflați în spațiul îngust al expertizei în date deschise și derivă din susținerea guvernamentală și din viziunea integratoare. Prima poate fi întărită prin promovare la nivel guvernamental, așa cum a reieșit din secțiunea dedicată notorietății. Cea de-a doua cunoaște provocarea proliferării site-urilor unde este difuzată informație dedicată unui singur domeniu. Data.gov.ro are nevoie de un set de acțiuni care să asigure cooperarea cu aceste platforme, iar capitolul final al acestui volum, recomandă metode dedicate de intervenție.

Din punct de vedere al disponibilității, apare nevoia de promovare a data.gov.ro și de monitorizare permanentă a gradului de utilizare. Un argument puternic este nevoia de comunicare a existenței și disponibilității portalului inclusiv la nivel politic. Acest lucru s-ar putea face prin implicarea directă a decidenților și prin includerea sistematică a domeniului în documentele de politici publice. Opinii pozitive și negativă concură către ideea că autoritățile publice locale trebuie să fie mai prezente în portal. Gradul de utilizare este considerat un indicator important de impact.

Discuția despre reutilizare a relevat mai degrabă probleme structurale ale portalului, care limitează numărul și calitatea utilizărilor. Participanții consideră că gradul de utilizare a crescut în principal prin prisma creștere numărului de seturi de date, deci o creștere inerentă. Principalele piedici sunt legate de valoarea economică scăzută a seturilor de date incluse în portal, dar și de sincope cu privire la actualizare sau de absența metadata care să le facă utilizabile.

METODA

Capitolul de față explorează rezultate ale celor două sondaje inițiate în cadrul proiectului de față. Ambele au fost colectate în perioada decembrie 2020-începutului lui ianuarie 2021. Contrastăm rezultate cu un sondaj realizat în 2017, după cum explicăm în continuare.

Primul sondaj din 2020 este un **sondaj adresat administrației publice locale**. Demersul a fost în acest sens mixt, implicând două moduri de abordare a potențialilor respondenți.

Pe de o parte, avem de a face cu un panel de respondenți ce au fost investigați și în trecut. Au fost trimise invitații de completare a chestionarului celor 388 de respondenți la sondajul despre date deschise inițiat de E & Y în 2017 la cererea SGG (Voicu, 2018)⁴ și care au indicat că pot fi recontactați pentru colectare ulterioară de date. Cei 388 au primit un link personalizat, netransmisibil. Acest lucru nu a afectat anonimitatea, prin modul de gestionare a sondajului neputându-se face legătura între respondent și răspunsurile sale.

Dintre cele 388 de adrese provenite din sondajul din 2017, aproximativ o sută de adrese s-au dovedit a fi nefuncționale, cel mai probabil aparținând unor funcționari care au părăsit fie administrația publică, fie organizația în care se aflau în urmă cu trei ani, fie (în puține cazuri) foloseau o adresă de email neinstituțională (de tipul @yahoo sau @gmail) și au renunțat la ea. Aceasta indică o fluctuație extrem de ridicată a personalului din administrația publică (ceva mai mult de un sfert). Dintre cei cu adrese valide, 17 au refuzat explicit să răspundă (au folosit opțiunea „opt out” prezentă în invitația la sondaj), 45 au accesat chestionarul dar nu au răspuns la nicio întrebare, fără a opta pentru un refuz explicit. Restul de 116 au oferit răspunsuri (complete sau parțiale) la sondaj.

Complementar, au fost trimise invitații de a răspunde la chestionar și de a difuza link-ul (nepersonalizat) în interiorul organizației lor către 75 de persoane din administrația publică, identificate de SGG ca puncte cheie în promovarea datelor deschise. Acest subeșantion, expus la același chestionar ca cel menționat mai sus, a furnizat 217 răspunsuri la chestionar și 173 de accesări fără a completa măcar un răspuns.

Al doilea sondaj a fost adresat mediului academic. Am selectat 12 universități, din care 10 de stat și 2 private, urmărit ca și criterii de stratificare gradul de specializare al universității, dimensiunea (ca număr de studenți) și dispersia geografică. Pentru cele 12 universități am colectat adresele de email ale corpului didactic, așa cum sunt ele disponibile pe site-urile oficiale ale departamentelor componente. A rezultat o listă de start de 6104 adrese. Este vorba în total de aproximativ 4900 de persoane (unele au avut 2 sau 3 adrese de email). Aproximativ 1000 de adrese s-au dovedit a fi nefuncționale, de această dată fiind vorba predominant de adrese neinstituționale. Dintre cei aproape 4000 de potențiali respondenți rămași, 128 au refuzat explicit să răspundă (opt out). 165 au accesat chestionarul fără a da vreun răspuns, iar 830 au răspuns la chestionar cel puțin parțial.

Chestionarul folosit a avut un modul comun care permite comparația între respondenții din 2017 și cei din 2020 și între cei din administrația publică și cei din mediul academic, precum și module care au fost aplicate doar în 2017, 2020, doar pentru mediul academic, respectiv doar pentru administrația publică.

⁴ Baza de date este disponibilă la <https://data.gov.ro/dataset/sipoca36>.

În cele ce urmează exploatăm pe rând aceste module, subliniind elementele comune și cele ce diferențiază. Studiem astfel, prin comparația cu 2017, trendurile observate și avem un potențial indicator privind impactul data.gov.ro în ultimii 3 ani.

De reținut că avem de a face cu eșantioane ne-probabilistice, astfel încât simpla comparare a procentelor sau mediilor observate poate conduce la concluzii distorsionate. Prezentăm totuși astfel de estimări univariate cu scop ilustrativ. Pentru a studia diferențele de interes, utilizăm analiză multivariată (analiză de regresie). Cu alte cuvinte, eliminăm diferențele induse de structura diferită a eșantioanelor și căutăm să aflăm ce s-a schimbat între 2017 și 2020 sau ce factori fac să fie diferențe în interiorul eșantioanelor din 2020.

Pentru întrebările comune folosite în 2017 și 2020 în cazul administrației publice locale, realizăm comparația și între cei ce au răspuns deopotrivă în 2017 și 2020. În câteva cazuri, folosim fixed effects models pentru a identifica factorii ce au condus la schimbări între cele două valuri ale sondajelor.

NOTORIETATEA

RATELE DE RĂSPUNS

Pattern-urile de răspuns la chestionar oferă indicații asupra cunoașterii domeniului datelor deschise. Lucrurile sunt simple: dacă nu știi nimic despre un domeniu, tinzi să nu răspunzi la întrebări. Caseta 3 oferă un exemplu de email primit de la unul dintre respondenți. Au fost destul de multe astfel de mesaje, venite mai ales dinspre mediul academic (cu precădere dintre cei din *humanities*), dar și dinspre administrația publică.

Caseta 3. Exemplu de mesaj de refuz de a răspunde

Buna ziua,

Eu acum aflu despre aceste "DATE DESCHISE" și nu le-am consultat, asa ca nu am cum sa raspund la chestionar.

Cu stima,

Fără îndoială, refuzul de a răspunde la chestionar nu e determinat exclusiv de necunoașterea domeniului. Pot fi nenumărate cauze care să conducă la rate de răspuns mici. Aceste cauze se regăsesc ca atare și în 2017 și în 2020, și la administrația publică și la mediul academic, adică la populația celor trei sondaje de care dispunem. E greu de spus dacă aceste cauze ce determină refuz de răspuns apar în egală măsură în cele trei eșantioane. Poate, cine știe, administrația publică a avut de răspuns în noiembrie-decembrie 2017 la atât de multe chestionare, încât nu a mai vrut să răspundă deloc... Nu avem de unde știi acest amănunt, dar am ales exemplul exagerat din fraza anterioară pentru a sublinia faptul că nu avem cum stăpâni complet variațiile ce pot apare în populațiile analizate.

Putem însă face asumptia că există cauze similare de a nu răspunde la chestionare la modul general și la chestionarele despre date deschise în particular.

Sondajul din 2017 s-a desfășurat în lunile de iarnă de la finalul anului, colectarea datelor s-a realizat prin aplicația Survey Monkey, iar unul dintre cei ce gestionau colectarea datelor a recurs la a suna șefi din organizațiile în care se aflau potențiali respondenți (cei selectați în eșantionul extras), cu scopul de a folosi autoritatea ierarhică pentru a crește ratele de răspuns. Este posibil ca aceasta să fi crescut ratele de răspuns. Sondajul dedicat administrației publice în 2020 a avut două tipuri de respondenți: pe de o parte cei ce indicaseră în 2017 că ar dori să participe la un nou sondaj, iar pe de altă parte pe cei către care punctele cheie identificate în organizații ale administrației publice au trimis invitația de a răspunde. În acest caz avem un alt tip de selecție pozitivă în eșantion: cei contactați aveau din start o probabilitate mai mare de a fi la curent cu datele deschise.

Tabelul 12 ilustrează situația răspunsurilor. Se observă că rata de răspuns la chestionar din total accesări este practic aceeași pentru administrația publică în 2017 și 2020. Această rată oferă o indicație în privința notorietății, în sensul discutat în paragrafele anterioare. Cei ce doar au accesat fără a completa sunt cei pentru care probabil subiectul este mai puțin familiar. Cele două rate sunt comparabile doar cu precauțiile deja luate, date de modelul diferit de solicitare a răspunsurilor. În plus, ne așteptăm ca în 2017 numărul celor spre care a mers invitația de a răspunde să fi fost mult mai mare decât în 2020, nu doar în termeni absoluți, dar și ca raport la total răspunsuri și accesări. Cu alte cuvinte, este probabil ca notorietatea problematicii datelor deschise să fi crescut în 2020 în raport cu 2017. Fără îndoială, această concluzie se bazează pe asumptii puternice, așa cum deja am explicat. Ea are nevoie de validare prin aducerea de informație complementară, așa cum facem în subcapitolul următor.

Tabelul 12. Număr de respondenți în funcție de poziționarea față de sondaje

	sondaj			Total
	Administrație Publică 2017	Administrație publică 2020	Academic 2020	
Doar au accesat chestionarul	663	218	165	1046
Au completat măcar un răspuns	1162	333	830	2325
Total	1825	551	995	3371
<i>Rate de răspuns raportat la total accesări</i>	<i>64%</i>	<i>60%</i>	<i>83%</i>	

*celulele includ număr de cazuri, exceptând ultimul rând din tabel.

Mai întâi însă, este util să observăm și pattern-urile de răspuns observate în cazul sondajului cu mediul academic. Au fost trimise invitații de a răspunde spre aproximativ 4000 de adrese valide, așa cum am explicat în debutul acestui capitol. Am obținut astfel o rată de accesare de aproximativ 25% și o rată de răspuns în total accesări de 83%. Cea din urmă reprezintă un indicator de notorietate ridicată printre cei ce au accesat chestionarul. De asemenea, rata de accesare de 25% este una bună în cazul unui astfel de sondaj, indicând o bună notorietate și interes față de domeniu.

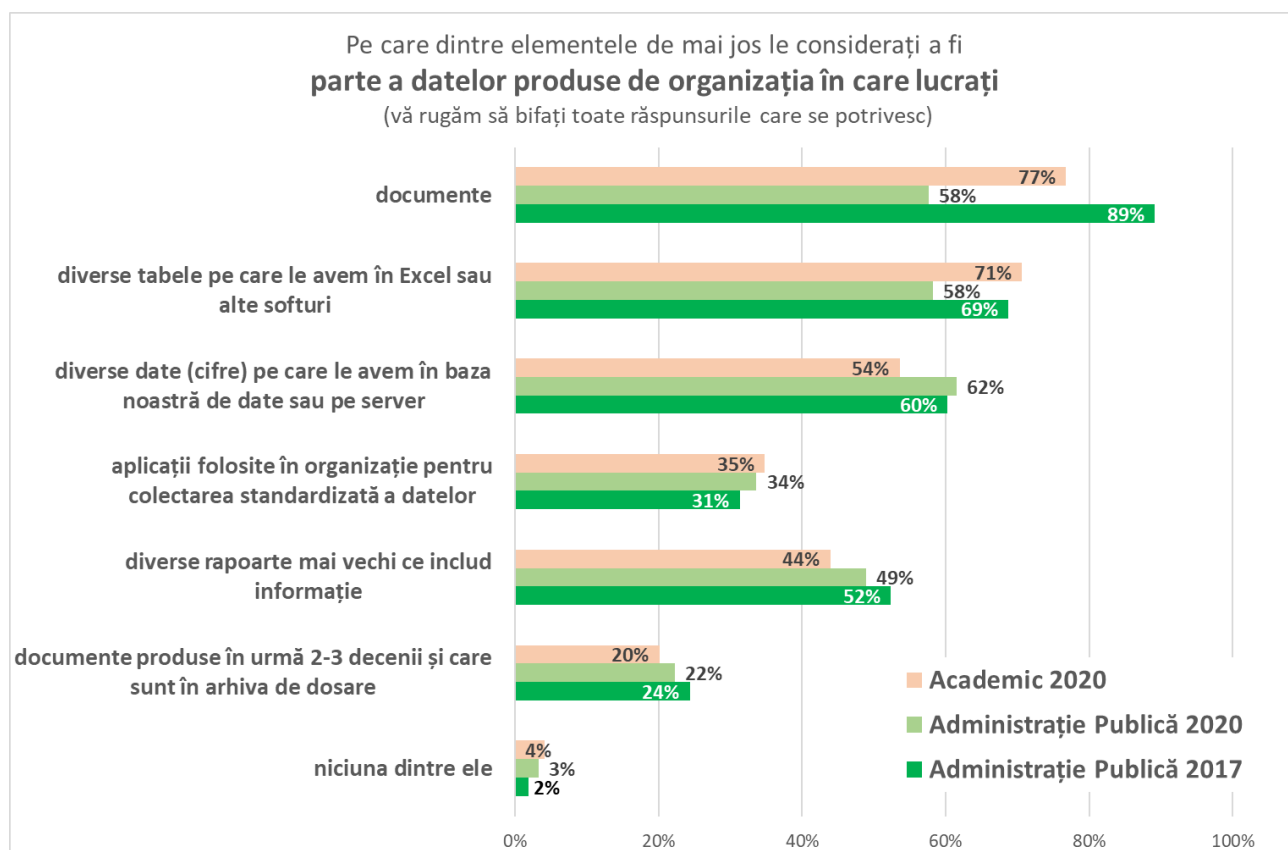
CUNOAȘTEREA

Cunoașterea despre ce sunt datele deschise (DD) este elementul cheie în toate analizele ce urmează. Ea condiționează practic opiniile despre data.gov.ro și chiar și utilizarea platformei, fiind o precondiție a implicării în deschiderea datelor în sine.

TIPURI DE DATE DESCHISE

O primă întrebare este ce ar putea reprezenta datele deschise în opinia respondenților. Am oferit câteva opțiuni de a alege respondenților, pentru a indica ce fel de date sunt produse în locul unde lucrează. Figura 5 ilustrează răspunsurile. Față de 2017 nu avem practic variații. Doar „documentele” definite ca date cunosc o scădere a frecvenței menționării. De asemenea, diferențele dintre mediul academic și administrația publică par a fi mai degrabă minore, putând fi induse de simple diferențe în structura eșantioanelor.

Figura 7. Opinii despre datele produse în organizația respondentului



*Cifrele reprezintă procente din respondenți ce au răspuns la chestionar, la nivelul fiecărui eșantion. Fiecare respondent a putut indica oricâte răspunsuri a dorit.

Este de observat că fiecare dintre elementele oferite spre alegerea respondenților sunt potențiale surse de date. Din acest punct de vedere, apare din nou nevoia de informare, care era vizibilă și în 2017. Probabil este utilă conștientizarea modurilor de utilizare a datelor și felul în care fiecare din sursele menționate devin

reutilizabile și mai ales utile în viața profesională cotidiană și pentru organizațiile din care respondenții fac parte.

Tabelul 6 adâncește analiza, realizând profilul celor ce au indicat fiecare din elementele propuse. Mai exact, folosind regresie logistică, putem spune care dintre caracteristicile individuale ale respondenților conduce la o probabilitate mai mare de a alege fiecare din acele elemente. Spre exemplu, prima linie din tabel are semnificația următoare: probabilitatea de a spune că „documentele” sunt parte potențială a datelor deschise este mai ridicată la cei chestionați în 2017, la cei care sunt între 35 și 39 de ani, la cei care vin din organizații mai mari de 60 de persoane, iar în eșantionul de universitari e vorba de cei din științele sociale sau din inginerie.

Tabelul 13. Categori de personal care au tendința de a indica mai des decât ceilalți anumite caracteristici ce fac datele să fie date

Pe care dintre elementele de mai jos le considerați a fi parte a datelor produse de organizația în care lucrați...?	Profilul celui ce alege mai frecvent varianta din prima coloană*
<ul style="list-style-type: none"> documente 	Administrație Publică – 2017 35-39 ani Mărime departament/organizație: peste 60 Academic: Științe sociale, Inginerie
<ul style="list-style-type: none"> diverse tabele pe care le avem în Excel sau alte softuri 	Academic 2020 Mărime departament/organizație: 11-30 Indicator ridicat de cunoaștere a DD Academic: are poziție de conducere Academic: nu este în Medicină
<ul style="list-style-type: none"> diverse date (cifre) pe care le avem în baza noastră de date sau pe server 	Admin publică 2020 55-59 de ani Indicator ridicat de cunoaștere a DD Academic: are poziție de conducere Academic: Științele naturii
<ul style="list-style-type: none"> aplicații folosite în organizație pentru colectarea standardizată a datelor 	Admin publică 2020 Academic 2020 55-59 de ani Mărime departament/organizație: 11-60 Indicator ridicat de cunoaștere a DD Academic: mai mulți ani de școală Academic: Științe Sociale, Științele naturii
<ul style="list-style-type: none"> diverse rapoarte mai vechi ce includ informație 	Admin publică 2020, (Admin publică 2017) 50-59 sau 35-39 de ani Indicator ridicat de cunoaștere a DD
<ul style="list-style-type: none"> documente produse în urmă cu 2-3 decenii și care sunt în arhiva de dosare 	Peste 30 de ani
<ul style="list-style-type: none"> niciuna dintre ele 	Academic 2020

*Profilul este realizat în fiecare caz folosind modele de regresie logistică, iar estimările sunt realizate ținând cont de stratificarea pe tipuri de organizații: ministere, autorități, consilii județene etc. Exemple de interpretare pot fi regăsite în text.

Investigând tabelul în ansamblul său, este de remarcat nivelul mai redus de menționare a aproape ficărui element în organizațiile mai mici (care de altfel au și mai puține date), dar și faptul că avem o cunoaștere mai puternică în eșantionul din 2020 al administrației publice. Acesta este de fapt rezultatul cel mai important, pentru că indică o creștere a notorietății în 2020 în comparație cu 2017.

Rezultatul ar putea fi însă doar un efect al creșterii cunoașterii la cei ce deja știau despre date deschise. Am studiat prin urmare schimbările doar în interiorul panelului celor ce au răspuns și în 2017 și în 2020⁵. Analiza relevă că, în fapt, în interiorul acestor aproximativ o sută de funcționari publici intervievați și în 2017 și în 2020, nu apar diferențe între cele două momente în care au fost colectate datele. Cu alte cuvinte, observăm în fapt o creștere a nivelului de cunoaștere la cei care au intrat în eșantion în 2020 și nu erau parte a eșantionului din 2017. În sine, acest lucru implică o creștere a cunoașterii la nivelul populației de funcționari publici.

CARACTERISTICI ALE DATELOR DESCHISE

Am chestionat de asemenea cunoașterea și altfel: la fel ca în 2017, am întrebat în ce măsură câteva elemente fac parte din modul de definire a datelor deschise. Tabelul 7 listează rezultatele. Am colorat cu albastru procentele care marchează diferențe consistente față de cifrele consemnate în 2017.

În primul rând, apar descreșteri ușoare dar generalizate, ale ponderii răspunsurilor „nu știu”, ceea ce întărește ipoteza unei cunoașteri superioare în 2020 comparativ cu 2017. Evident, toate acestea trebuie interpretate pentru moment sub precauția eșantioanelor neprobabilistice și cu o structură diferită.

Tabelul 14. Reprezentări despre caracteristicile datelor deschise („În ce măsură fiecare dintre următoarele elemente credeți că se aplică datelor deschise (open data)”)

	Administrație publică				Mediul academic 2020	
	2017		2020			
	Nu știu	În mare sau foarte mare măsură	Nu știu	În mare sau foarte mare măsură	Nu știu	În mare sau foarte mare măsură
sunt date brute în format tabelar	10%	47%	9%	60%	10%	64%
sunt produse de organizații private	23%	14%	16%	20%	18%	34%
sunt protejate de drepturi de autor	23%	21%	14%	22%	13%	47%
sunt lipsite de valoare	20%	9%	12%	5%	17%	11%
au valoare comercială	22%	22%	17%	26%	22%	30%
pot fi și documente scanate ca imagine sau .pdf	8%	55%	7%	45%	10%	60%
sunt disponibile online	6%	70%	5%	86%	5%	82%
pot fi reutilizate (pot fi folosite și în alte scopuri decât cel inițial, precum aplicațiile)	14%	57%	9%	81%	11%	75%
Au caracter personal și confidențial	12%	22%	8%	14%	10%	18%
oricine poate folosi oricând datele deschise	10%	65%	7%	87%	9%	77%
Datele deschise și informațiile publice sunt una și același lucru	15%	43%	10%	44%	17%	34%

*Răspunsurile posibile au fost: 1=în foarte mică măsură ... 4=în foarte mare măsură, respectiv „nu știu”; refuzurile de a răspunde sunt ignorate în calculul procentelor de mai sus.

⁵ Am folosit în acest sens aceleași modele de regresie logistică, dar am restrâns eșantioanele la cei din administrația publică care au răspuns și în 2020 și în 2017. Rezultatul comentat în continuare se referă la efectul momentului culegerii datelor: în 2017 sau în 2020.

Dincolo de aceasta, tendința generală este de a avea creștere a reprezentării că sunt date în format tabelar, protejate de drepturi de autor, disponibile online și pentru toți, cu capacitate de reutilizare. Astfel de trenduri sunt greu de interpretat, deoarece ele relevă deopotrivă cunoaștere și necunoaștere.

Pentru a analiza mai ușor schimbarea între 2017 și 2020, precum și factorii asociați cu cunoașterea mai bună a DD, la fel ca în studiul realizat pe datele din 2017 (Voicu, 2018), am construit un indicator de cunoaștere, al cărui mod de construcție este prezentat în Caseta 4. Indicatorul a fost deja folosit în mod tacit în analizele din subcapitolul precedent și urmează a fi exploatat și în restul acestui volum.

Caseta 4. Modul de construcție a indicatorului de cunoaștere a datelor deschise

Pentru a utiliza în mod sintetic datele din Tabelul 14, a fost generat un indicator de cunoaștere al conținutului acestora. Indicatorul este construit ca o notă școlară. Au fost luați în considerare 10 dintre cei 11 itemi din tabel. Fiecare item a fost notat cu un punct sau cu zero puncte, după corespondența cu definiția datelor deschise. Itemul neluat în calcul este „datele deschise sunt lipsite de valoare”. Motivul rezidă în sensul ambiguu al termenului „valoare”, care poate trimite fie la faptul că nu au valoare comercială, fie la valoarea lor în sine, fie ea și din punct de vedere al cunoașterii generate. Pentru cei zece itemi rămași, s-a marcat un punct dacă:

- s-a răspuns cu „în foarte mare măsură” sau „în mare măsură” la definirea datelor deschise prin:
 - sunt date brute în format tabelar;
 - oricine poate folosi oricând datele deschise;
 - sunt disponibile online;
 - pot fi reutilizate (pot fi folosite și în alte scopuri decât cel inițial, precum aplicațiile),
- s-a răspuns cu „în foarte mică măsură” sau „în mică măsură” la definirea datelor deschise prin:
 - sunt produse de organizații private;
 - sunt protejate de drepturi de autor;
 - au valoare comercială;
 - pot fi și documente scanate ca imagine sau .pdf;
 - Au caracter personal și confidențial;
 - Datele deschise și informațiile publice sunt una și același lucru.

Cu alte cuvinte, fiecare respondent a primit o notă de la 0 la 10 la cunoașterea datelor deschise.[...]. O astfel de notare este una laxă, deoarece am permis respondenților să se îndoiască ușor de greutatea afirmației. O notare mai dură ar însemna să luăm în considerare doar variantele extreme („în foarte mare măsură”, respectiv „în foarte mică măsură”). Rezultatul îl constituie un indicator de cunoaștere dur.

*conținutul casetei este extras din Voicu (2018, p. 15-16)

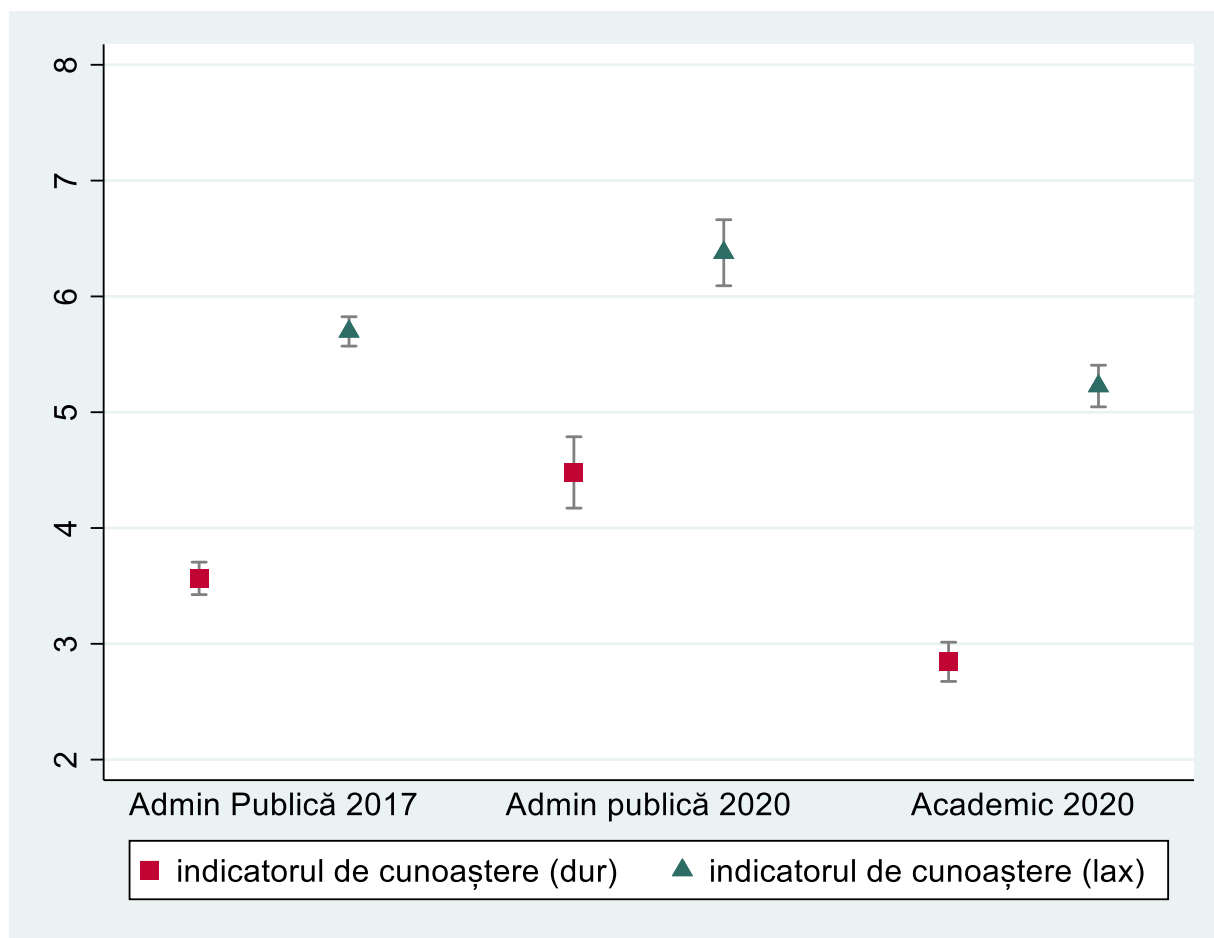
Tabelul 15 ilustrează distribuția celor trei eșantioane conform acestui indicator de cunoaștere, în cele două variante în care a fost calculat. Figura 8 prezintă mediile celor doi indicatori și intervalul de încredere (locul unde s-ar afla media dacă am extrage 95 de eșantioane similare). În cazul administrației publice, se remarcă scorurile mai mari observate în 2010. În cazul mediului academic, se remarcă scorul sistematic mai mic. Diferențele între administrație publică și mediul academic, precum și cele între 2020 și 2017 rămân importante și după ce eliminăm efectele vârstei, dimensiunii organizației, educației etc., așa cum am realizat și în cazul celorlalte analize. Indicatorul de cunoaștere este mai ridicat pentru cei care au între 35 și 39 de ani,

apoi descrește cu vârsta, dar cei mai puțin informați rămân cei sub 30 de ani. Dimensiunea organizației contează de asemenea, cunoașterea fiind mai redusă în organizațiile/departamentele mai mici (sub 10 persoane).

Tabelul 15. Structura eșantioanelor conform indicatorilor de cunoaștere calculați

		sondaj		
		Admin Publică 2017	Admin publică 2020	Academic 2020
indicatorul de cunoaștere (lax)	0-4	29%	23%	38%
	5-7	49%	37%	43%
	8-10	22%	40%	19%
	Total	100%	100%	100%
indicatorul de cunoaștere (dur)	0-4	66%	51%	77%
	5-7	28%	36%	21%
	8-10	6%	14%	2%
	Total	100%	100%	100%

Figura 8. Medii ale indicatorului de cunoaștere (cu estimări ale intervalelor de încredere 95%)



*descrierea celor doi indicatori de cunoaștere este realizată în text. Fiecare respondent putea primi maxim 10 puncte și minim 0 pentru oricare dintre cei doi indicatori.

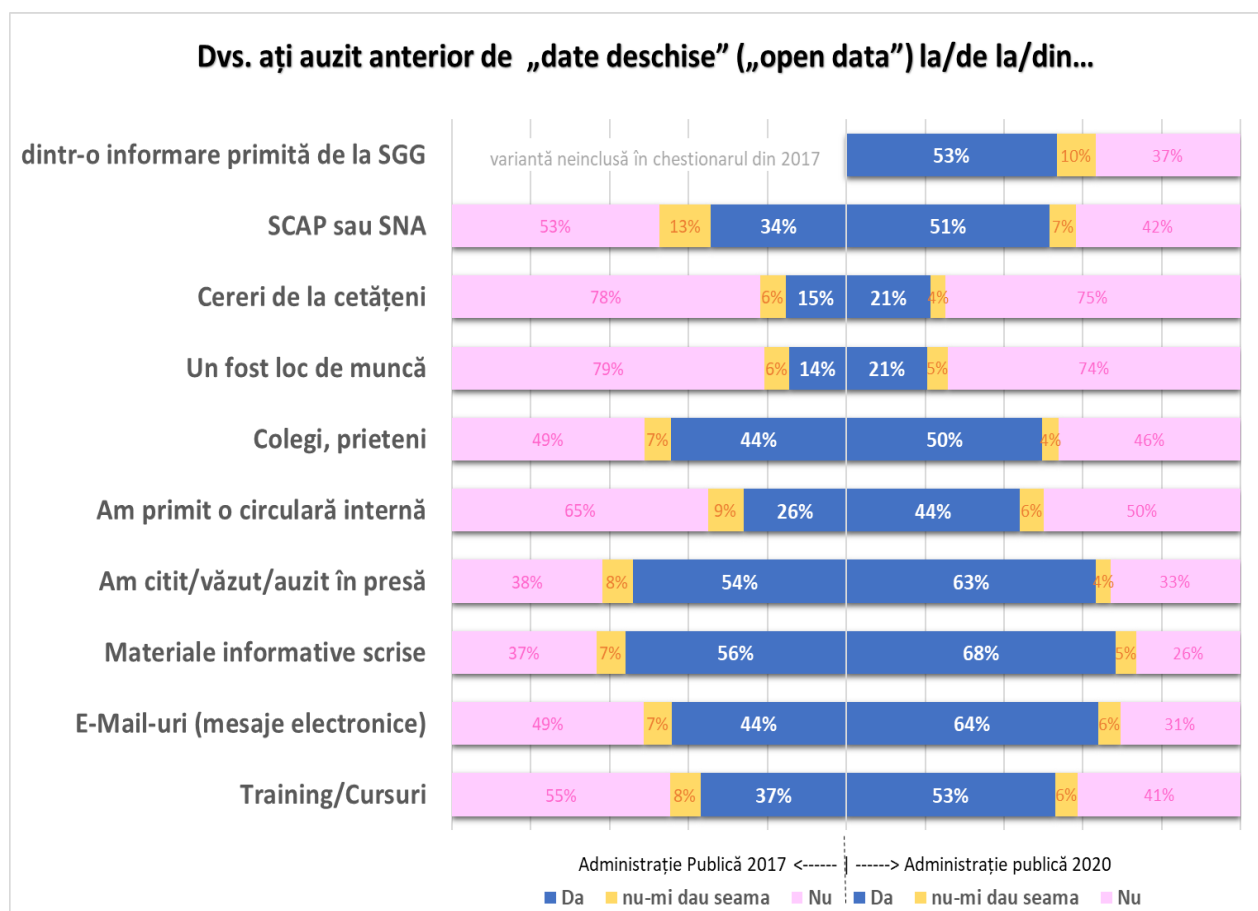
Dacă pentru administrația publică luăm în calcul doar pe cei care au răspuns și în 2017 și în 2020, constatăm că indicatorul de cunoaștere practic nu variază între cele două momente în timp (Figura 8).

SURSE ALE CUNOAȘTERII

Am chestionat respondenții din eșantionul de funcționari publici și în privința locului de unde au aflat despre date deschise (Figura 9). Comparația cu 2017 este grăitoare la nivel de distribuție a frecvențelor: majoritatea surselor sunt menționate de mult mai mulți respondenți decât în urmă cu trei ani, vorbind de la sine despre circulația informației despre date deschise. Portofoliile respondenților continuă a fi variate, jumătate dintre respondenții din 2020 indicând minim 4 surse din care au aflat despre date deschise.

Comunicarea directă, în scris sau electronic, precum și știrile din media sunt sursele prevalente de informare, urmate de traininguri și colegi. Între sursele scrise/electronice, prevenirile din SCAP (Strategia pentru consolidarea administrației publice 2014–2020) sau cele din SNA (Strategia Națională Anticorupție) sunt menționate de jumătate dintre respondenții din 2020, același lucru petrecându-se cu informările de la Secretariatul General al Guvernului.

Figura 9. Surse de informare despre open data



*SCAP este Strategia pentru consolidarea administrației publice 2014–2020, SNA reprezintă Strategia Națională Anticorupție.

ROMÂNIA VS RESTUL LUMII: REPREZENTĂRI SUBIECTIVE

Și în 2017 și în 2020, sondajele au inclus o întrebare evaluativă despre publicarea de date deschise în România comparativ cu restul Europei. Tabelul 16 prezintă rezultatele. Coloanele din stânga compară eșantioanele complete din 2017 și 2020. Coloanele din dreapta sunt mai precise, prin faptul că oferă o imagine a schimbării în timp pentru exact aceeași respondenți din administrația publică ce au răspuns și în 2017 și în 2020.

Analiza rezultatelor din panel indică faptul că practic distribuția răspunsurilor nu s-a schimbat între 2017 și 2020. Cu alte cuvinte, privind acum către coloanele din stânga, avem o majoritate care consideră că datele deschise din România sunt mai puține ca în alte țări UE. Pesimismul în această privință este mai pregnant în mediul academic.

Tabelul 16. Evaluarea subiectivă a publicării de date deschise în România și în UE

Ați spune că România publică date deschise ...	sondaj			Panel	
	Admin Publică 2017	Admin publică 2020	Academic 2020	2017	2020
nu a răspuns	0%	11%	1%	0%	7%
nu știu	45%	26%	26%	30%	25%
mai puțin ca în alte țări UE	41%	49%	62%	57%	53%
la fel ca în alte țări UE	13%	13%	10%	13%	15%
mai mult ca în alte țări UE	1%	1%	0%	0%	0%
Total	100%	100%	100%	100%	100%

*Panelul include cei aproximativ 100 de respondenți care au răspuns și în 2017 și 2020.

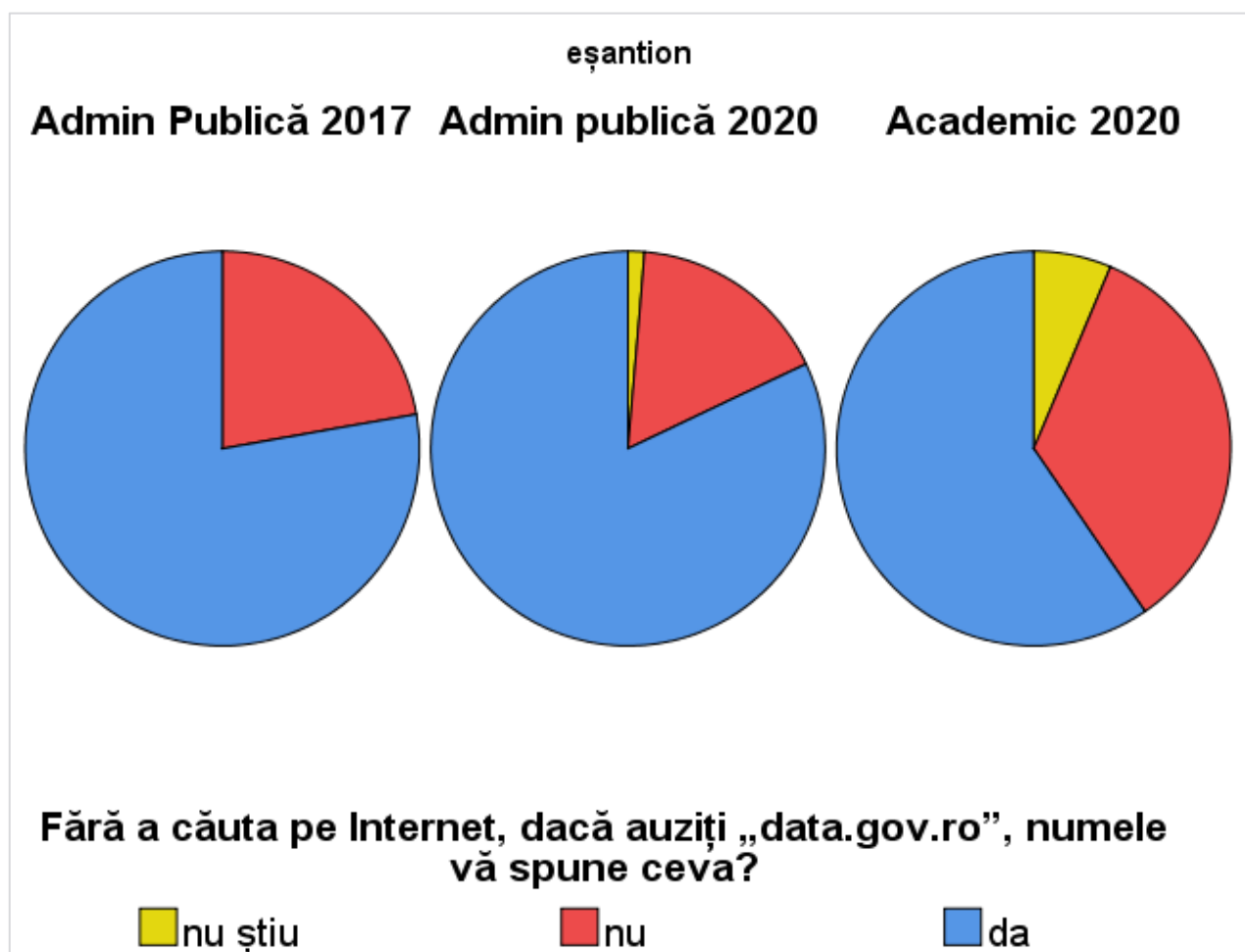
DATA.GOV.RO

Notorietatea site-ului data.gov.ro continuă să fie ridicată în 2020. În 2017, 79% dintre respondenți declarau că au auzit de platforma guvernamentală, cifra fiind practic neschimbată în 2020 (diferența nu este „semnificativă”). În schimb, în mediul academic, cunoașterea data.gov.ro scade la sub două treimi (60%). Dacă eliminăm efectele de structură a eșantioanelor, diferențele dintre 2017 și 2020 rămân practic nule în ce privește administrația publică, în timp ce mediul academic are o șansă de 3 ori mai mică de a fi informat despre data.gov.ro în comparație cu funcționarii publici⁶.

⁶ Odds ratio. Acest paragraf, precum și următorul, prezintă rezultate din analiza de regresie logistică, dublată de fixed effects models.

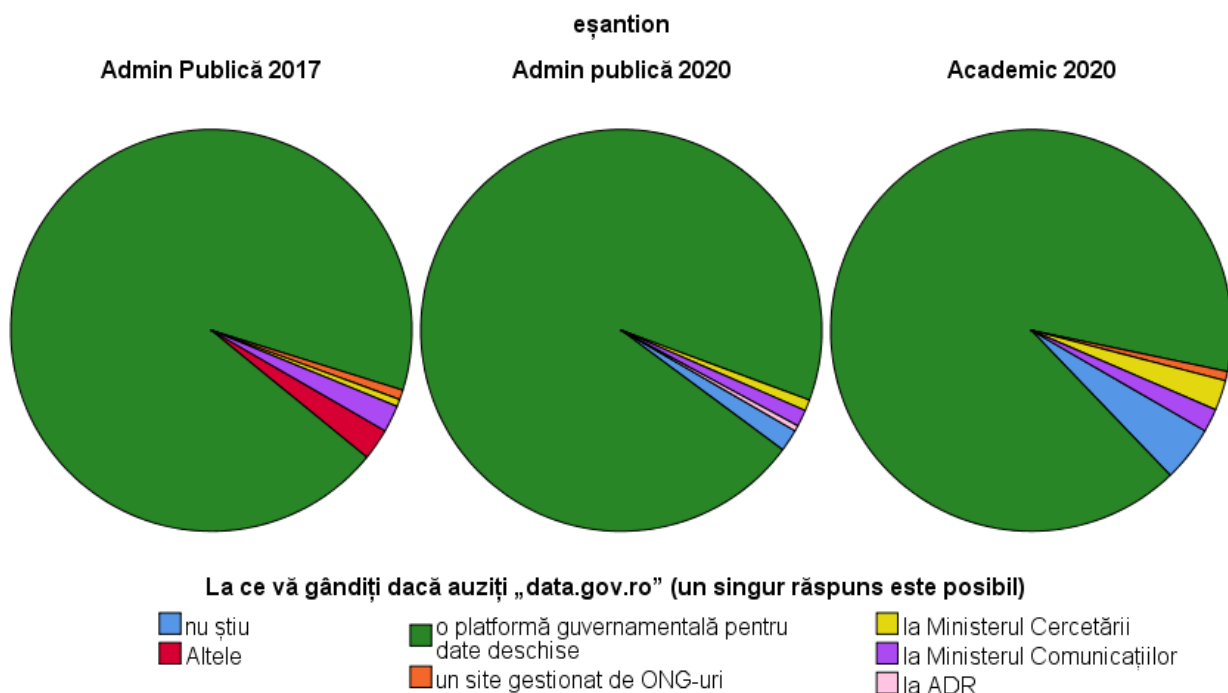
În plus, se observă o notorietate declarată mai redusă în rândul celor între 30 și 45 de ani, precum și în rândul celor ce au cunoștințe mai puține despre datele deschise (cu alte cuvinte, cunoașterea datelor deschise este asociată cu informarea despre data.gov.ro). Între universitari, cunoașterea data.gov.ro scade odată cu gradul didactic, de la profesor la conferențiar și lector (cele două grade sunt similare din punct de vedere al cunoașterii data.gov.ro) și apoi la asistenți și preparatori. În plus, cei din științele naturii au un grad de cunoaștere mai ridicat, iar cei din medicină și artă au o cunoaștere mai redusă.

Figura 10. Notorietatea data.gov.ro în cele trei eșantioane



Cunoașterea data.gov.ro nu este una întâmplătoare: în toate cele trei eșantioane, peste 90% dintre respondenți o desemnează corect drept o platformă guvernamentală, estimările fiind de 94% în 2017, 96% pentru funcționarii publici chestionați în 2020, respectiv 91% pentru mediul academic. (Figura 11)

Figura 11. Localizarea data.gov.ro



LEGITIMITATEA

MODURI DE DESCHIDERE A DATELOR: MEDIUL ACADEMIC

Chestionarul adresat mediului academic a inclus un pachet de itemi ce urmărea să observe moduri de a deschide datele. Răspunsurile sunt ilustrate în Tabelul 17. Pentru 50% dintre respondenți, datele deschise ar trebui să devină publice și accesul să fie gratuit, 36% cred că ele trebuie să fie publicate pe pagina de Internet a organizației, iar un sfert menționează nevoia de a acorda o licență liberă. Doar 9% spun că nu doresc să publice date deschise, iar 11% nu știu cum se poate face acest lucru. Cu alte cuvinte, avem de a face cu o legitimitate medie a publicării de date deschise (prima întrebare), dar responsabilitatea pentru distribuire ar trebui să aparțină altcuiva. Procentul redus legat de licența liberă este probabil legat de necunoaștere.

Diferențierile în interiorul eșantionului sunt practic inexistente⁷. Doar femeile au șanse⁸ de patru ori mai mari de a spune că nu doresc publicarea deschisă a datelor și de trei ori mai mici că nu doresc asigurarea accesului gratuit la date. În plus, cei cu funcții de conducere au șanse de 9 ori mai mici de a dori publicarea deschisă a datelor pe pagina de Internet a departamentului sau universității. În cazul ultimilor, efectul observat întărește supoziția că refuzul de a publica date pe pagina proprie este legată în principal de costurile de timp și resursă umană legate de acest demers.

⁷ Paragraful acesta explorează rezultate ale analizelor de regresie logistică. Variabilele independente folosite sunt: gender, vârstă, ani de școală absolviți, grad didactic, a deține funcție de conducere, domeniul în care activează (artă, științe sociale, științe naturii, științe exacte, inginerie, medicină, economie, drept, științe umaniste), indicatorul de cunoaștere a datelor deschise).

⁸ Este vorba de odds ratios. Adică de raportul dintre opiniile pozitive și cele negative, ca la cotele de la pariuri.

Tabelul 17. Modalități considerate legitime de a deschide datele: mediul academic, 2020

Dintre următoarele, ce măsuri ați luat sau credeți ar trebui să luați pentru a publica datele produse de dvs. ca date deschise?	nu este bifat	nu a răspuns la pachet	este bifat	Total
asigurarea accesului gratuit tuturor celor interesați	23%	27%	50%	100%
publicarea pe pagina de Internet a organizației	37%	27%	36%	100%
asocierea unei licențe libere	48%	27%	25%	100%
nu știu	73%	16%	11%	100%
nu doresc să public date deschise	73%	18%	9%	100%

În context, să notăm că jumătate din eșantion declară că fie personal fie prin echipele de care aparține a depozit deja date în arhive de date, 25% nu știu dacă au făcut-o, iar un alt sfert spune că nu a făcut-o vreodată. Este de remarcat și faptul că nu există vreo asociere între opiniile din Tabelul 17 și experiența de a depune date în arhive de date⁹.

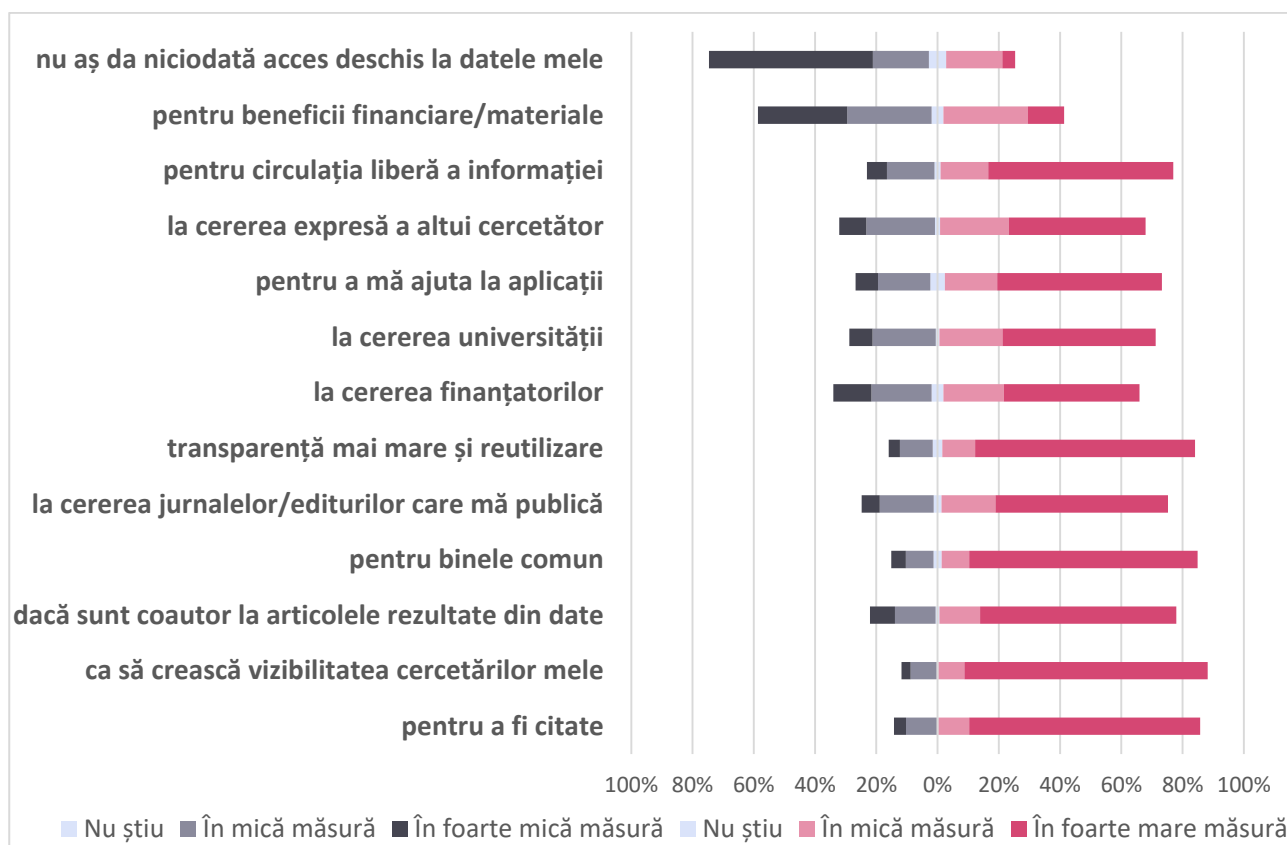
MEDIUL ACADEMIC: CONDIȚII DE DESCHIDERE A DATELOR

În chestionarul adresat mediului academic am inclus două module de întrebări referitoare la motivele care ar împiedica sau accelera deschiderea datelor. Figura 12 ilustrează condițiile în care este confortabil pentru respondenți ca să deschidă datele produse de ei. Se observă o prevalență a motivațiilor legate de libera circulație (transparență, binele comun) și beneficii academice (citare, vizibilitate). Motivațiile extrinseci, legate de constrângere sunt mai slabe dar acționează suficient de puternic. Majoritatea respinge ideea beneficiilor materiale și a neoferirii de date deschise.

Figura 13 pune în evidență motivele de reținere în a deschide datele. De această dată, imaginea care se înfățișează este mai complexă. Cea mai mare parte a motivelor de stopare a deschiderii datelor atrag majorități ce le resping, sugerând că nu constituie bariere reale în calea deschiderii datelor. Excepția vine din două direcții diferite. Pe de o parte este vorba de acea perioadă de embargo tipică în mediul academic, în care cel ce produce datele este cel ce are primatul în analiza acestora. Ea vine împreună cu potențiale probleme de copyright. Pe de altă parte este o chestiune de resurse, în speță de timp pentru deschiderea datelor.

⁹ De asemenea, e vorba de rezultate din analiză multivariată (regresie logistică).

Figura 12. Mediul academic: În ce condiții vă simțiți confortabil să lăsați libere datele produse de dvs.



Cele 13 motivații de deschidere și 15 motive de reținere la deschidere pot fi grupate în 3 dimensiuni¹⁰. Primul factor este unul de acceptare a deschiderii datelor și de acord cu factorii motivaționali. Al doilea factor grupează majoritatea motivelor de reținere, al treilea factor respinge motivele de transparență și bine comun, și accentuează beneficiile instrumentale derivate din deschidere și cele de protecție a informației conținute.

Avem prin urmare un factor motivațional general, unul de respingere generalizată și un al treilea de respingere instrumentală.

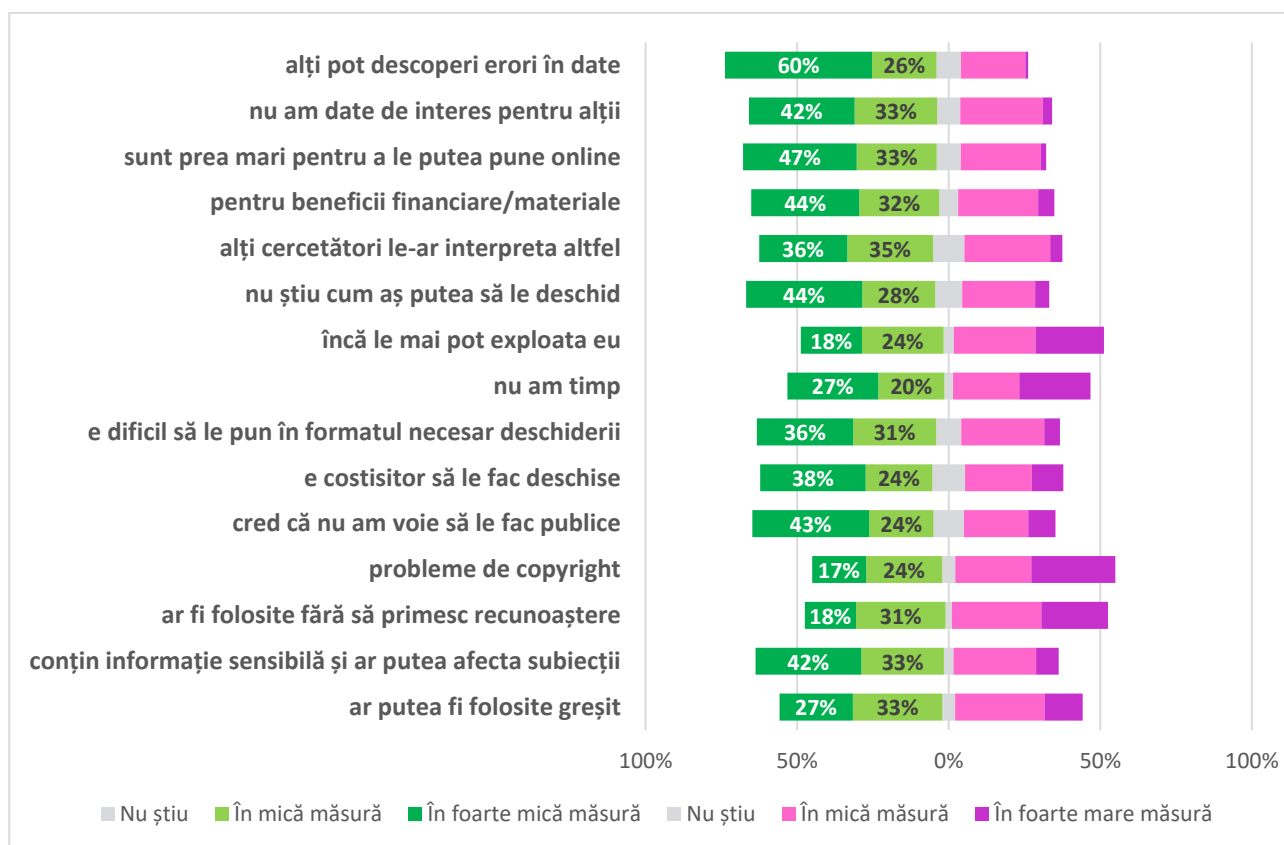
Motivele de respingere instrumentală¹¹ sunt asociate cu creșterea în vârstă, dar scad la profesori și conferențieri, unde crește motivația generală de a deschide date. De asemenea, motivația instrumentală este mai accentuată în Arte și Drept.

În rest, nu sunt diferențe mari de la un grup de status la altul, cu o singură excepție: indicatorul de cunoaștere crește puternic motivația generalizată de a deschide date și scade respingerea instrumentală. Din nou, de dovedește utilitatea cunoașterii în a accelera deschiderea datelor.

¹⁰ Prin analiză factorială. Cei trei factori ortogonali explică 38% din variația totală, toate comunalitățile sunt peste 0.2, metoda de extracție a fost Maximum likelihood, iar rotația Equamax.

¹¹ Acest paragraf exploatează rezultate din analiza de regresie a celor trei factori.

Figura 13. Mediul academic: Ce probleme/rețineri aveți legate de a deschide datele produse de dvs.



ACCEPTAREA REUTILIZĂRII

Legitimitatea deschiderii datelor este dată în bună măsură de acceptarea ideii că datele respective pot fi utilizate de către terțe părți. Sondajul cu administrația publică din 2020 a inclus o baterie de itemi ce testează legitimitatea utilizării datelor în diferite scopuri.

Tabelul 18 ilustrează răspunsurile. Opiniile dominante sunt favorabile utilizării datelor în practic orice scop, în spiritul esenței deschiderii datelor. Acceptarea (re)utilizării este însă mai puternică în ce privește scopuri legate de administrația publică sau academice, și ceva mai redusă în ce privește utilizare în programe ale partidelor politice. Reținerea din urmă este firească în atmosfera de neîncredere în sistemul politic prezentă în România.

Susținerea pentru reutilizare în oricare dintre scopurile din Tabelul 18 este asociată cu o probabilitate ridicată de a susține reutilizarea în oricare din celelalte scopuri. Cu alte cuvinte, pentru un respondent oarecare, avem de a face de fapt cu o orientare generală de susține sau nu reutilizarea datelor, care se reflectă în susținere specifică pentru fiecare scop în parte din cele menționate. Variația în interiorul eșantionului nu este foarte mare. Se observă însă o tendință clară de a legitima reutilizarea datelor dacă ai cunoaștere mai bună despre datele deschise, precum și la cei care au peste 50 de ani.

Tabelul 18. Reprezentări ale administrației publice în ce privește legitimitatea (re)utilizării datelor deschise în funcție de scopul (re)utilizării

În ce măsură sunteți de acord ca datele produse în departamentul sau organizația dvs. să fie reutilizate ...	Nu răspund	Nu știu	În foarte mică măsură	În mică măsură	În mare măsură	În foarte mare măsură	Total
• În politici publice	6%	4%	8%	7%	35%	40%	100%
• În decizii punctuale ale administrației publice	7%	3%	7%	8%	34%	41%	100%
• În mediul de business	7%	6%	13%	19%	29%	26%	100%
• Academic: cercetare și sau predare	3%	2%	1%	3%	28%	63%	100%
• În analize ale ONG-urilor	8%	6%	12%	14%	27%	33%	100%
• În programe ale partidelor politice	11%	8%	33%	15%	14%	19%	100%
• De către jurnaliști	10%	7%	21%	15%	23%	24%	100%

Tabelul 19. Deschiderea spre utilizarea datelor deschise

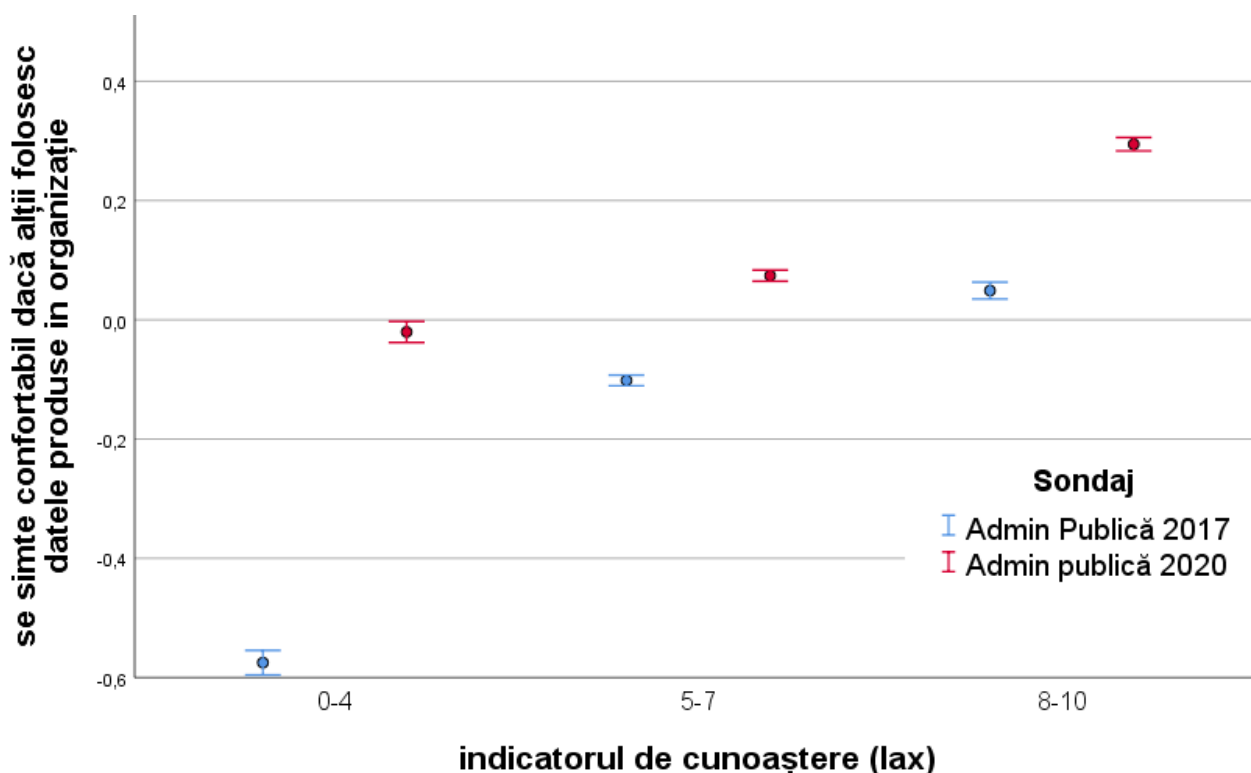
	Ponderea celor ce au răspuns pozitiv*	
	Admin Publică 2017	Admin publică 2020
Vă simțiți confortabil să lăsați colegi din același departament să folosească datele dvs.	79%	85%
Vă simțiți confortabil să lăsați colegi din ALT departament să folosească datele dvs.	64%	79%
Vă simțiți confortabil să lăsați colegi din alte organizații publice să folosească datele dvs.	42%	66%
Vă simțiți confortabil să lăsați pe oricine să folosească datele dvs.	28%	40%

*Variantele de răspuns au fost scalate de la 1 (în foarte mică măsură) la 4 (în foarte mare măsură). Procentele din tabel reprezintă ponderea celor ce au răspuns "în foarte mare măsură" sau "în mare măsură".

Folosind răspunsurile la întrebările despre acceptarea utilizării datelor deschise descrise în Tabelul 19, am construit un indicator de deschidere la către reutilizare. Calculat ca scor factorial¹², indicatorul variază de la -2 (respingere completă a ideii de reutilizare) la +2 (acceptare completă).

Figura 14 descrie variația indicatorului în 2017 și 2020. Observația simplă este legată de o creștere a deschiderii spre utilizare și reutilizare, aceasta menținându-se și după ce eliminăm efecte legate de structura eșantioanelor¹³ și observându-se o dependență de gradul de cunoașterea datelor deschise: cei ce au mai multe cunoștințe despre date deschise sunt mai deschiși spre utilizarea datelor lor de către terți.

Figura 14. Cunoaștere despre date și deschidere către reutilizare



Cifrele reprezentate grafic constituie medii ale indicatorului de reutilizare în funcție de nivelele indicatorului (lax) de cunoaștere. Ambii indicatori sunt descriși în text. Indicatorul de acceptare a reutilizării datelor variază de la -2 (respingere cvasitotală a reutilizării) la +2 (acceptare totală). Punctele roșii și cele albastre sunt estimări ale mediilor. Dreptele din jurul lor constituie intervalele în care s-ar plasa aceste puncte în 95 din 100 de eșantioane (în ipoteza că eșantioanele din 2017 și 2020 ar fi fost probabilistice).

Este de notat și faptul că efectul cunoașterii era mai puternic în 2017. Acest lucru în sine este pozitiv, indicând faptul că între 2017 și 2020 avem pe de o parte o schimbare în ce privește deschiderea către utilizare și reutilizare, iar această schimbare este mai puternică pentru cei ce au mai puține cunoștințe despre date deschise. Pe ansamblu, avem de a face cu un efect de contagiune și multiplicare a suportului pentru reutilizarea datelor.

Utilizarea vine cu costuri asociate reduse, în opinia respondenților din administrația publică, iar aceste opinii nu s-au modificat în 2020 în comparație cu 2017, după cum o arată Tabelul 20. Foarte puțini sunt cei care spun că ar trebui să existe o plată pentru accesarea datelor. Suportul pentru ca datele să fie complet deschise

¹² KMO=.764, variația explicată de factorul unic extras prin Maximum Likelihood: 71%, toate comunitățile sunt peste .5.

¹³ Prin analiză de regresie.

este din acest punct de vedere remarcabil. În schimb, la nivelul mediului academic, susținerea pentru deschiderea gratuită este mai redusă, exceptând deschiderea în interiorul mediului academic.

Tabelul 20. Costul utilizării datelor deschise

	Ponderea celor ce au răspuns pozitiv*		
	Admin Publică 2017	Admin publică 2020	Academic 2020
<i>Eșantion</i> →			
Organizațiile publice ar trebui să plătească pentru utilizarea datelor produse de organizația dvs.	13%	11%	36%
Organizațiile private ar trebui să plătească pentru utilizarea datelor produse de organizația dvs.	23%	33%	65%
ONG-urile ar trebui să plătească pentru utilizarea datelor produse de organizația dvs.	16%	17%	34%
Datele produse în organizația dvs. ar putea fi folosite <u>fără plată</u> în scopuri academice (educație, cercetare)	71%	87%	88%

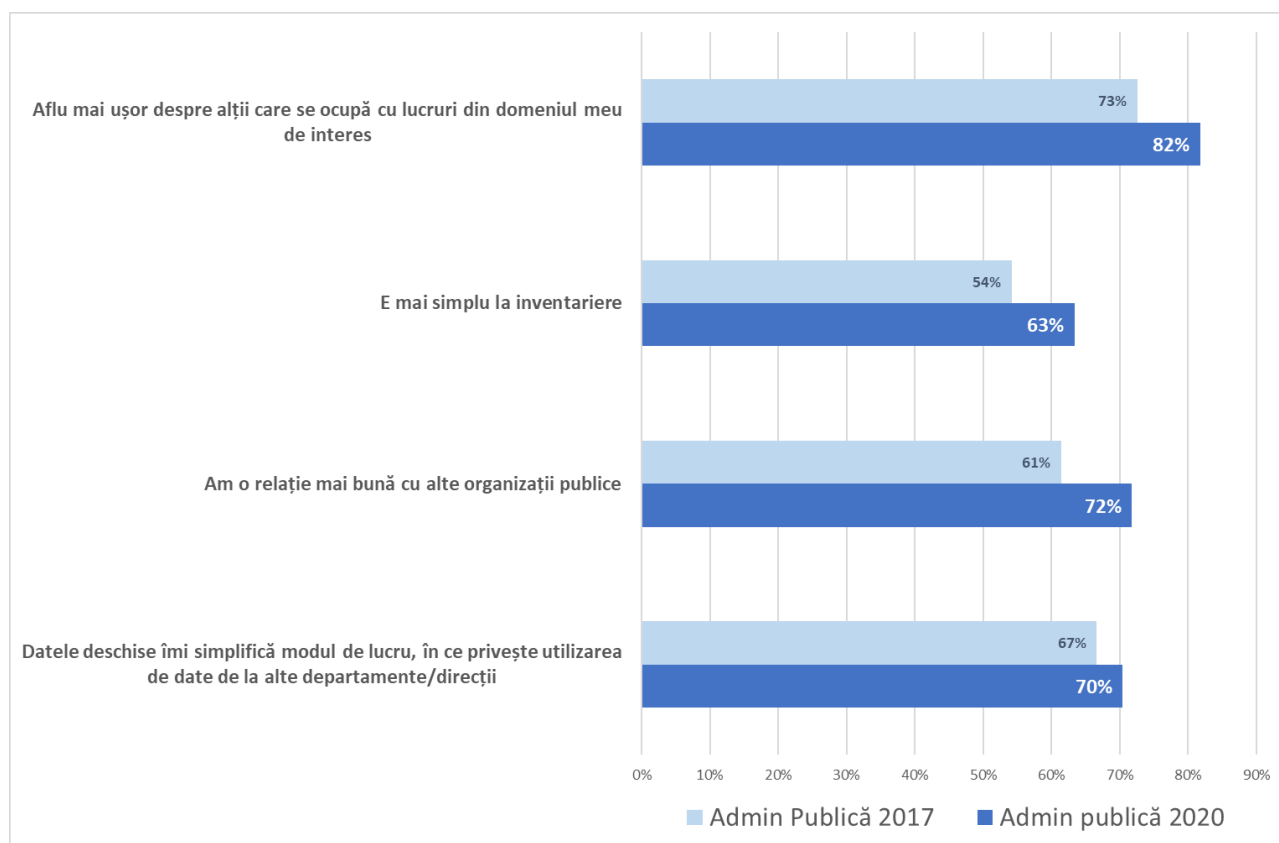
*Variantele de răspuns au fost scalate de la 1 (în foarte mica măsură) la 4 (în foarte mare măsură). Procentele din tabel reprezintă ponderea celor ce au răspuns “în foarte mare măsură” sau “în mare măsură”.

OPINII DESPRE UTILITATEA DATELOR DESCHISE

Ca și în 2017, și în 2020 chestionarul relevă afirmarea faptului că existența datelor deschise aduce beneficii (Figura 22). Fără îndoială, există o probabilitate destul de mare ca simpla acceptare de a răspunde la chestionar să implice un efect de dezirabilitate conducând la a tinde să răspunzi pozitiv la întrebările despre utilitatea datelor deschise. Tocmai de aceea este importantă comparația între cele două momente în timp în care am realizat măsurarea acestor opinii.

Din acest punct de vedere, graficul relevă creșteri suficient de mari ale opiniilor pozitive despre cele patru caracteristici luate în considerare, pentru ca aceste creșteri să indice o modificare a poziționării administrației publice. Modificarea se păstrează și după ce elimină efecte ale structurii eșantioanelor și indică o creștere a ponderii celor care apreciază utilitatea datelor deschise.

Figura 15. Opinii despre beneficiile deschiderii datelor



*Variantele de răspuns au fost scalate de la 1 (în foarte mica măsură) la 4 (în foarte mare măsură). Procentele din figură reprezintă ponderea celor ce au răspuns “în foarte mare măsură” sau “în mare măsură”.

DISPONIBILITATEA

CONSTRÂNGERI ȘI CATALIZATORI AI DESCHIDERII DATELOR

Tabelul 21 descrie reprezentările celor trei eșantioane despre factorii ce ar putea fi implicați în deschiderea datelor. Comparând seria de cifre pentru administrația publică în 2017 și 2020 se observă creșteri ușoare, dar aflate în marja de eroare a estimării, exceptând o creștere a reprezentărilor despre importanța acțiunilor SGG și a expunerii publice. Practic, deși putem avea un simplu efect al cunoașterii inițiatorului sondajului de către respondenți, avem un indiciu asupra aprecierii eforturilor SGG de popularizare. În acest sens este de remarcat și scăderea numărului de răspunsuri „nu știu” între 2017 și 2020.

Cifrele pot fi interpretate și prin raportare la nivelul general al importanței acordate celorlalți factori. Din această perspectivă, accentul cade în mod mult mai explicit în 2020 pe elemente intrinseci ale procesului, legate de calitatea datelor și instrumentele folosite pentru deschidere. Cu alte cuvinte avem de a face cu o focalizare mai puternică pe procesul de deschidere a datelor și pe mărirea acestora, și mai puțin pe elementele de natură politică și presiune societală.

Tabelul 21. Opinii despre factorii ce pot contribui la deschiderea datelor

Mai jos sunt listate mai multe elemente ce pot fi importante în producerea și diseminarea datelor (deschise).

Pentru fiecare vă rugăm să evaluați pe o scală de la 0 (total negativ) la 10 (total pozitiv), ce impact au în acest moment, în România, asupra datelor deschise:

Eșantion → Întrebări ↓	Admin Publică 2017		Admin publică 2020		Academic 2020	
	% nu știu	media	% nu știu	media	% nu știu	media
cadrul normativ (juridic) existent	24%	6,8	21%	6,9	33%	5,8
Atitudinea și agenda managementului superior	21%	6,7	16%	6,7	23%	5,9
Abilitățile (priceperea) celor din sistem	18%	6,9	13%	7,2	18%	6,1
timpul disponibil	21%	6,3	14%	6,5	21%	6,2
Presiunea societății civile	26%	6,5	19%	6,4	23%	5,8
Presiunea comunității universitare / de cercetare	15%	7,0
Calitatea datelor	22%	7,0	13%	7,5	14%	7,3
Instrumentul de publicare	22%	7,0	13%	7,5	13%	7,2
Presiunea politică	35%	5,8	26%	5,9	32%	4,8
Costurile	33%	6,1	24%	6,2	21%	6,5
Nevoia de anonimizare	33%	6,3	22%	6,9	23%	6,1
Capacitatea de a le exporta în formatul dorit	26%	6,8	13%	7,3	20%	6,5
posibilitatea de a valorifica datele prin vânzare	45%	5,4	45%	5,2	28%	5,5
expunerea publică	28%	6,5	21%	7,3	20%	6,5
Acțiunile SGG	37%	6,5	26%	7,4	44%	5,7
Pregătirea publicului	27%	6,5	.	.	23%	6,0
Apartenența României la UE	26%	7,5	17%	8,0	21%	7,7
Opiniile funcționarilor publici despre datele deschise	30%	6,0	21%	6,5	32%	5,2
Opiniile politicienilor despre datele deschise	40%	5,6	32%	5,9	34%	4,7
Schimbările de conducere din instituțiile publice	38%	5,5	26%	6,1	33%	5,5
Limitările de software care împiedică exportul de date	33%	6,5
Capacitatea de a exporta date a software-ului existent	.	.	19%	7,4	23%	6,4
capacitatea de reutilizare de la nivelul societății / comunităților	.	.	17%	7,4	.	.

* celulele goale corespund elementelor neincluse în sondajul respectiv.

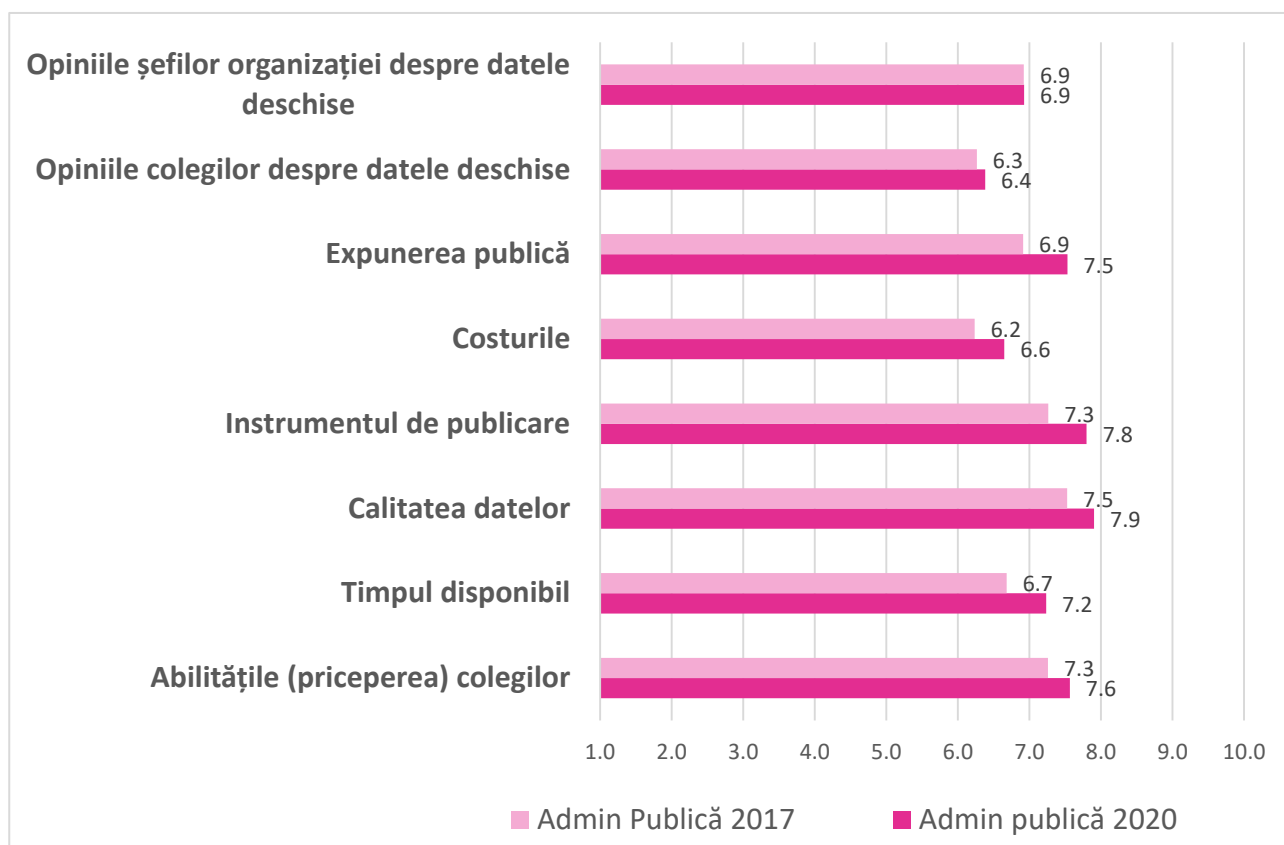
Efectul este încă și mai vizibil în cazul mediului academic, care are însă și particularitatea de a fi mai puțin preocupat cu abilitățile de a mânui date și cu prezența software-ului, concentrarea fiind efectiv pe calitatea datelor și capacitatea de publicare.

Se adaugă elementul de prestigiu, dat de importanța percepută a apartenenței la UE.

Vorbim prin urmare despre un mix al profesionalizării deschiderii datelor și despre nevoia de a o face ca afirmare a apartenenței la un spațiu internațional contemporan ce are asociat un prestigiu ridicat. Concentrarea pe acești factori poate constitui o ancoră de creștere a disponibilității datelor, precum și un indiciu al impactului proceselor ce au fost manifeste în ultimii ani.

Figura 16. Opinii despre factorii ce pot facilita deschiderea datelor în organizația de apartenență

Să ne întoarcem la organizația unde lucrați. Ce impact credeți că au următoarele asupra datelor deschise.



*cifrele reprezintă medii observate la nivelul celor două eșantioane pe scala de la 1 la 10, unde 10 indică impact total pozitiv, iar 1 – impact total negativ.

Figura 16 aduce informației suplimentară, de astă dată doar cu referire la administrația publică și concentrându-se pe organizația din care face parte respondentul. Din nou, ponderea non-răspunsurilor scade la aproape jumătate, iar mediile indică concentrarea pe aceleași elemente legate de calitatea datelor și abilitățile de mănuiere a acestora.

COSTURI, GRATUITATE ȘI RESTRICȚIONĂRI ALE ACCESULUI

În 2020, opinia despre necesitatea gratuității datelor deschise este și mai accentuată în comparație cu 2020. Ea este însoțită de o deschidere aproape universală și de conștientizarea beneficiilor imediate pentru organizație. La nivelul administrației publice apare controversa asupra nevoii de a răsplăti deschiderea datelor, în timp ce mediul academic solicită explicit o astfel de recunoaștere.

Tabelul 22. Opinii despre eventuala monetizare sau deschidere restricționată a datelor publice

	eșantion		
	Admin Publică 2017	Admin publică 2020	Academic 2020
Este incorect să facem datele publice, pentru că le vor folosi și cei care nu au plătit pentru ele	21%	14%	18%
Datele deschise ne permit să beneficiem gratuit de expertiza celor ce le folosesc și le analizează	71%	84%	85%
Trecerea la date deschise este inevitabilă	66%	85%	79%
Organizațiile care publică date deschise ar trebui recompensate	35%	45%	67%
Datele deschise conduc la economie de timp: fiind deja deschise ne scutesc de la a răspunde solicitărilor de a le furniza	73%	84%	85%
Datele ar trebui deschise doar către cei ce oferă și ei date	.	16%	23%
Datele ar trebui deschise doar către cei din același domeniu sau din sectoare conexe	.	17%	28%

* celulele goale corespund întrebărilor neincluse în sondajul respectiv.

Pentru datele colectate în 2020 am construit doi indicatori ce descriu orientările (latente) aflate în spatele răspunsurilor la întrebările din Tabelul 22 Figura 16¹⁴. Primul indicator este de respingere a ideii de limitare a accesului la date, corelând puternic negativ cu primul și ultimii doi itemi din Figura 16, și având corelații pozitive moderate cu restul indicatorilor. Al doilea factor este cel al reprezentărilor despre interesul pentru beneficiile derivate din deschiderea datelor, și are corelații pozitive cu toți itemii, moderate cu cei patru ce se referă explicit la beneficii și mai reduse cu cei care indică respingerea deschiderii necondiționate de plată sau participare la procesul deschiderii. Analizând cei doi factori¹⁵, nu apar diferențe semnificative între mediul academic și administrația publică, se observă obișnuita dependență de cunoașterea datelor deschise, o respingere a deschiderii nelimitate pentru cei ce vin din organizații între 11 și 30 de angajați și mai ales între 31 și 60 de angajați. Interesul pentru beneficii (și monetizare) este ridicat la cei mai tineri, apoi scade la cei între 40 și 50 de ani, pentru a crește iar ușor la grupele de vârstă de peste 50 de ani.

¹⁴ Am folosit în acest sens analiza factorială, extracția fiind realizată prin Maximum Likelihood, iar rotația fiind de tip Equamax. Factorii ortogonali explică 45% din varianța totală, KMO=.711, testul grohotișului indică clar doi factori, iar toate comunalitățile sunt mai mari de 0.218.

¹⁵ Modele de regresie cu estimarea robustă a erorilor standard.

RESPONSABILITĂȚI ȘI PROCEDURI ÎN DESCHIDEREA ȘI GESTIUNEA DATELOR

Dincolo de orice, disponibilitatea datelor depinde de oameni care să pună în mișcare datele, pregătindu-le pentru *upload* și urcându-le în platforme precum *data.gov.ro*, arhive de date dedicate, sau pe site-ul propriu. Figura 17 subliniază confuzia ce continuă să persiste la nivelul responsabilităților privind persoanele ce au responsabilități în deschiderea datelor. Indiferent de eșantionul analizat, opiniile sunt împărțite între a plasa deschiderea datelor în sarcina departamentelor de comunicare și a celor IT, iar jumătate dintre respondenți cred că publicarea datelor deschise constituie o sarcină suplimentară consumatoare de timp. De aici nevoia de a fi furnizate modele practice, ușor de înțeles și implementat în deschiderea și publicarea datelor.

Figura 17. Responsabilități în deschiderea datelor

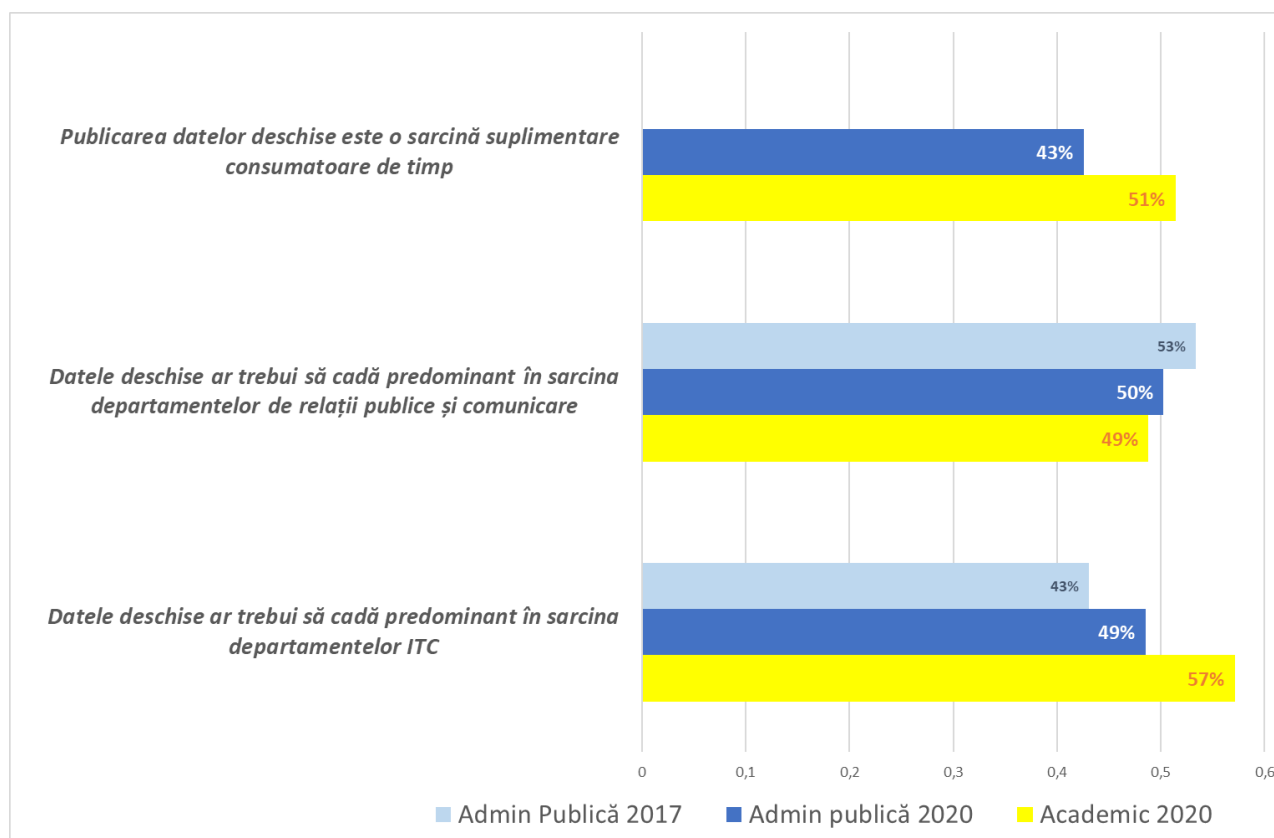


Figura 18 relevă mai în amănunt nevoia de modele. Ponderea răspunsurilor „nu știu” la întrebările despre procedurile („În organizația dvs., există un regulament intern în care sunt prevăzute proceduri legate de date deschise”) și gestiunea datelor („În organizația dvs., există un Inventar al seturilor de date”) este covârșitoare prin comparație cu cei care spun că există astfel de proceduri sau inventarieri ale seturilor de date. Situația, neschimbată față de 2017, relevă un deficit de cunoaștere. Totodată, se observă relativ ușor cum în 2020 cei care știu despre prezența regulamentelor interne ce guvernează datele deschise le și folosesc. Acest din urmă element subliniază faptul că prezența unui „model” (în cazul de față regulamentul intern) este de natură a produce consecințe. Chiar și dacă cei ce răspund de fapt nu ar ști despre regulamentul intern, răspunsul lor de aliniere la standardele (necunoscute ale) regulamentului indică capacitatea procedurilor scrise de a produce consecințe, fie ele și limitate.

Figura 18. Proceduri de gestiune a datelor în administrația publică



Tabelul 23 evidențiază în continuare un nivel de confuzie în ce privește guvernarea datelor deschise. Numărul răspunsurilor „nu știu” este în continuare foarte mare. Totuși numărul acestor indicații despre lipsa de cunoaștere asupra temelor propuse sunt în scădere în 2020 în comparație cu 2017, indicând un priceps de creștere a notorietății datelor deschise și a gândirii despre guvernanta acestora.

Altfel, ansamblul răspunsurilor zugrăvește un tablou destul de clar conturat. Respondenții nu prezintă o opinie majoritară despre cooperarea dintre organizații, așteaptă o abordare top-down ceea ce legitimează rolul unui organism central dedicat deschiderii datelor (în cazul de față, SGG), apare respingerea ideii descentralizării, și nu apar temeri legate de GDPR. Nevoia de a avea o abordare top-down și siguranța că nu apar probleme legate de viața privată a cetățenilor sunt elementele care atrag majorități ale răspunsurilor mai mari în 2020 decât în 2017.

Tabelul 23. Opinii despre guvernarea datelor deschise

		Admin Publică 2017	Admin publică 2020
Există bună coordonare între varii organizații publice	Nu știu	24%	10%
	În foarte mică măsură	17%	16%
	În mică măsură	24%	26%
	În mare măsură	24%	26%
	În foarte mare măsură	11%	21%
	<i>Total</i>	<i>100%</i>	<i>100%</i>
O abordare de sus în jos (top-down) este potrivită în domeniul datelor deschise	Nu știu	28%	21%
	În foarte mică măsură	8%	4%
	În mică măsură	14%	11%
	În mare măsură	31%	37%
	În foarte mare măsură	19%	27%
	<i>Total</i>	<i>100%</i>	<i>100%</i>
Poate fi lăsată la latitudinea autorităților locale	Nu știu	23%	17%
	În foarte mică măsură	29%	36%
	În mică măsură	25%	26%
	În mare măsură	17%	17%
	În foarte mare măsură	5%	4%
	<i>Total</i>	<i>100%</i>	<i>100%</i>
Publicarea de date deschise pune probleme legate de viața privată a cetățenilor	Nu știu	26%	14%
	În foarte mică măsură	30%	47%
	În mică măsură	24%	25%
	În mare măsură	13%	9%
	În foarte mare măsură	7%	5%
	<i>Total</i>	<i>100%</i>	<i>100%</i>

NEVOI DE FORMARE

În sondajele din 2020, cei chestionați au fost solicitați să spună în ce măsură au nevoi de formare. Tabelul 24 relevă răspunsurile. Atât printre funcționarii publici, cât și în mediul academic, mai bine de trei sferturi dintre respondenți indică nevoia de formare. Este de remarcat și faptul că nu sunt diferențieri majore în interiorul celor două eșantioane: doar asistenții universitari resimt o nevoie mai acută de formare decât restul.

Tabelul 24. Dvs. credeți că ați avea nevoie de cursuri certificate în deschiderea datelor?

	survey	
	Admin publică 2020	Academic 2020
nu	24%	24%
da, pentru utilizarea datelor	69%	73%
da, pentru publicarea datelor	70%	69%

*suma procentelor poate da peste 100%, pentru că un respondent putea alege ambele variante "da".

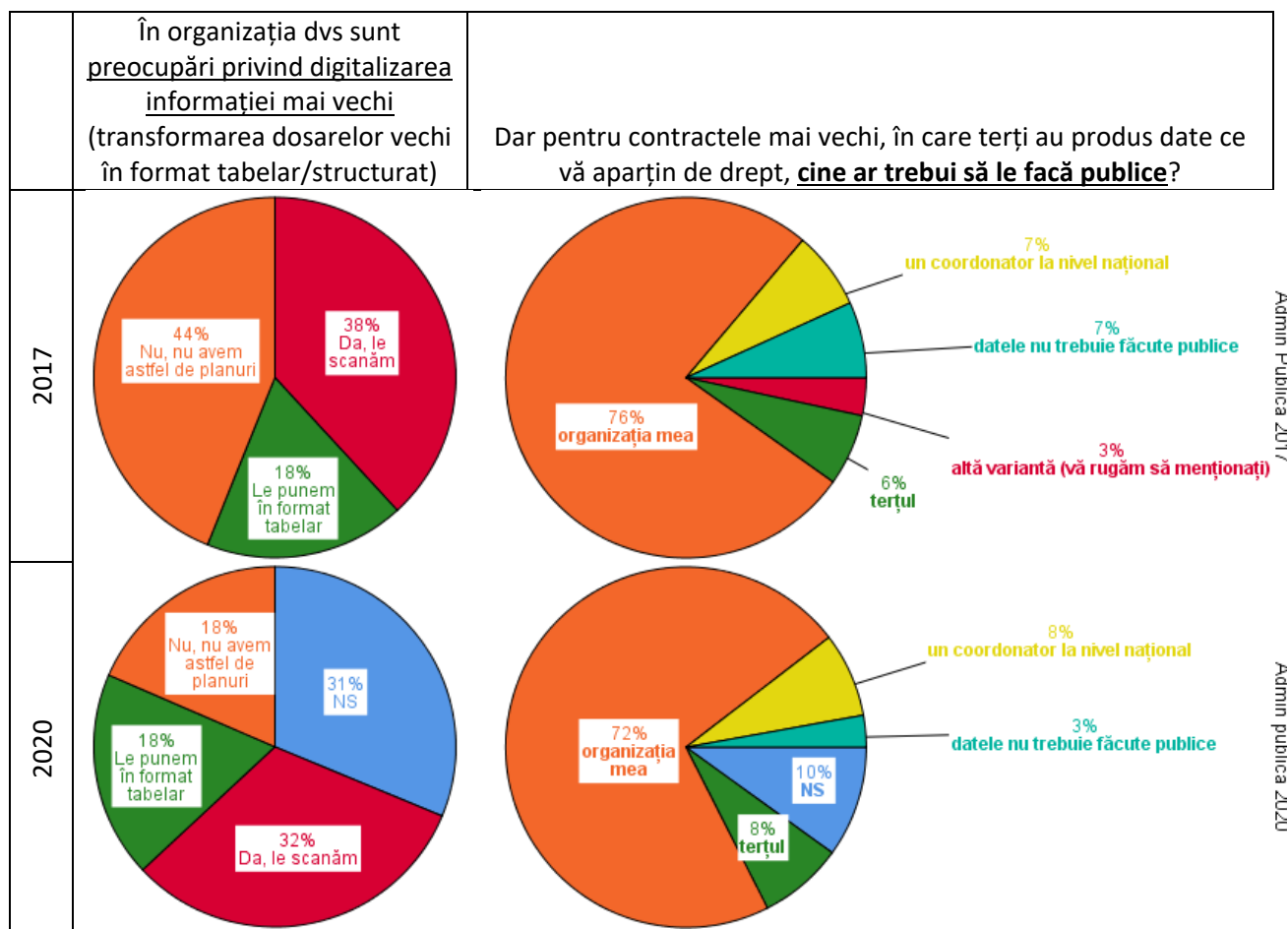
RECUPERAREA SETURILOR DE DATE VECI ȘI DATELE PRODUSE DE TERȚI

O bună parte din datele aflate în arhivele administrației publice sunt date vechi, disponibile în fișiere disparate în computere sau pe hard-disk-uri externe sau chiar pe hârtie. A aduce aceste date în circulație prezintă în principiu două tipuri de provocări: prima este cea a utilității. Stereotipul românesc, inclusiv în mediul academic, este că datele „vechi” sunt inutile, iar ceea ce ne trebuie sunt date noi, actualizate de preferință ieri, dacă nu cumva chiar mâine. Numai că datele din administrația publică nu sunt perisabile, iar datele vechi permit evaluări ale impactului politicilor, analize ale trendurilor, fundamentarea politicilor curente, etc. Fără îndoială, ele au capacitatea de a fi exploatate de către mediul economic, academic și de către societate în ansamblul său, dar pentru a pune accentul pe beneficiile imediate, este important să observăm mai ales avantajele deschiderii pentru administrația publică în sine.

A doua dificultate majoră e dată costurile de timp implicate pentru a aduce datele în format analizabil. Aceasta implică adesea un volum de muncă mare, ce poate fi perceput ca piedică în calea recuperării.

Chestionarele adresate administrației publice au inclus și în 2017 și în 2020 întrebări ce se referă la aceste aspecte. Pe de o parte am dorit să observăm dacă există preocupări pentru a aduce datele mai vechi în spațiul public, iar pe de alta dacă ar fi legitim ca terțe părți să fie implicate în procesul de a aduce datele la format analizabil.

Figura 19. Preocupări privind deschiderea datelor mai vechi



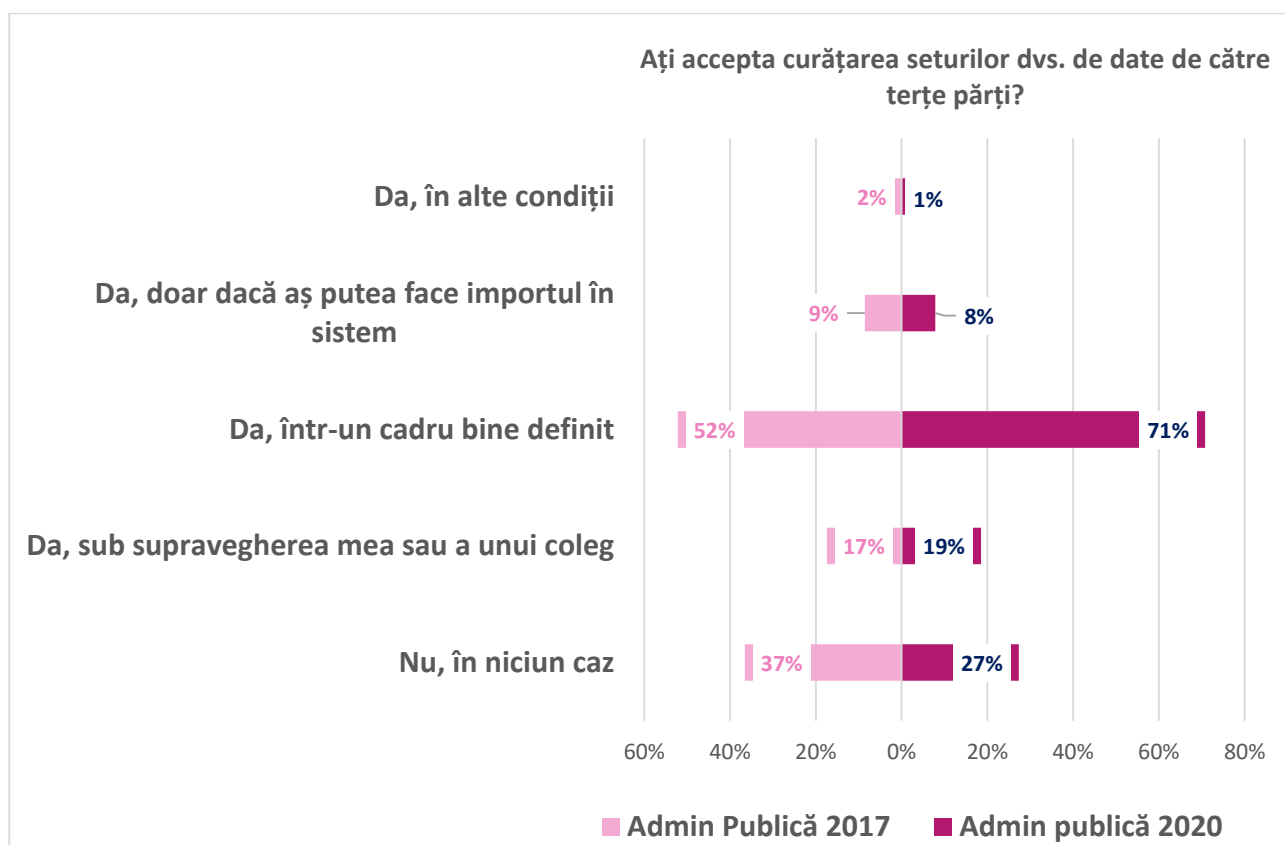
*valul 2017 nu a inclus varianta de răspuns „nu știu”

Figura 19 ilustrează interesul primar pentru punerea în valoare a datelor mai vechi. Chiar dacă cifrele nu sunt perfect comparabile, date fiind diferențele din modul de întrebare, se remarcă ponderea probabil staționară a celor care nu au preocupări de aducere a datelor mai vechi în circuitul datelor deschise (sau nu știu despre ele), și faptul că și ponderea celor ce spun că le aduc în spațiul public, fie în format tabelar (18% în 2020) sau măcar prin scanare (32% în 2020) este de asemenea aproape aceeași ca în 2017. Figura 19

În privința aducerii în spațiul public a datelor produse în trecut de către terți pentru administrația publică, există o opinie majoritară puternică, neschimbată între cele două momente în timp, că organizația proprietară (i.e. administrația publică) este cea care ar trebui să le aducă în spațiul public.

Opinia se menține și în privința datelor produse în prezent de către terți la cererea organizațiilor publice: 61% dintre respondenții din administrația publică spun în 2020 că organizația lor ar trebui să publice datele, 21% declară că este responsabilitatea terților, 8% invocă un coordonator la nivel național, 3% se opun deschiderii unor astfel de date, iar 8% nu știu ce să spună. Cifrele erau similare și în 2017, cu observația că varianta „nu știu” nu fusese inclusă în chestionar.

Figura 20. Opinii despre folosirea terților în procesul de deschidere a datelor



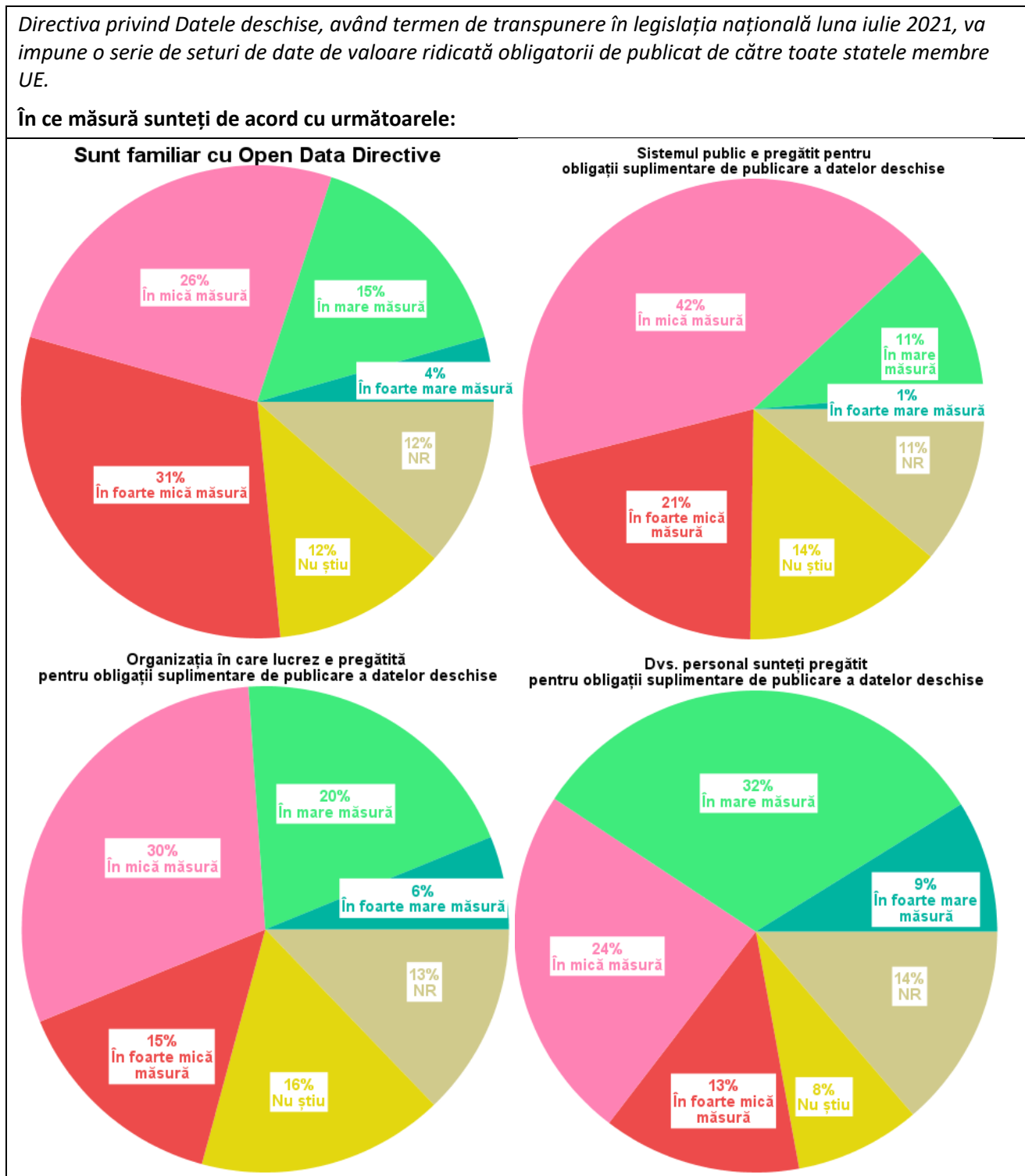
*procentele nu sunt cumulative, fiecare respondent a putut alege mai multe variante de răspuns

Lucrurile se schimbă în privește procesul de pregătire pentru publicare. Majoritatea respondenților acceptă implicarea terților, iar creșterea față de 2017 este notabilă (Figura 20).

OPEN DATA DIRECTIVE

Directiva privind Datele Deschise urmează a produce efecte de la jumătatea lui 2021. Am testat în ce măsură este cunoscută și cum se raportează administrația publică la posibilitatea de a o implementa.

Figura 21, Opinii ale administrației publice legate de implementarea Open Data Directive



Datele indică pe de o parte necunoașterea care domină reacțiile la Open Data Directive. Pe de altă parte relevă în majorități puternice scepticismul în ce privește posibilitățile sistemului public de a implementa directiva. Un scepticism ușor mai moderat privește organizația din care face parte respondentul și pe respondent în sine, dar pe ansamblu asistăm la puține opinii favorabile (feliile verzi din Figura 21)

INCIDENȚA UTILIZĂRII DE DATE PRINTRE RESPONDENȚI

ADMINISTRAȚIA PUBLICĂ

Sondajul cu administrația publică din 2020 a inclus o baterie de întrebări ce solicitau respondentului să specifice dacă folosește seturi de date (1) în scop profesional; (2) preluate de pe data.gov.ro; (3) preluate din alte surse; sau (4) produse de propria organizație. Este vorba de orice fel de date, nu doar de date deschise.

Tabelul 25 descrie răspunsurile. Să notăm pentru început că jumătate dintre respondenții ce au citit bateria de întrebări nu au răspuns deloc, neprecizând nici dacă folosesc, nici că nu folosesc date de vreun fel. Există o probabilitate mare ca acești respondenți să nu folosească deloc seturi de date sau să nu conștientizeze că ar folosi date. În absența unui mod de a distinge cele două situații, putem concluziona că non-răspunsul reprezintă mai degrabă o declarație non-utilizare a seturilor de date. Rămânem cu 41% dintre respondenți care declară că folosesc date (în cazul a 48% este incert, 11% nu folosesc deloc). 27% au folosit în scop profesional, 5% au preluat date de pe data.gov.ro, 17% din alte surse, iar 22% folosesc date produse de propria organizație.

Tabelul 25. Utilizarea de date la nivelul eșantionului de angajați din administrația publică din 2020

Dvs. ați folosit vreodată seturi de date ...	Neselectat (nu folosește)	nu a răspuns la pachetul de itemi	Selectat (folosește)	Total
în scop profesional	25%	48%	27%	100%
preluate de pe data.gov.ro	47%	48%	5%	100%
preluate din alte surse	35%	48%	17%	100%
produse de propria dvs organizație	30%	48%	22%	100%
niciuna din cele de mai sus	41%	48%	11%	100%

Am căutat să aflăm cine sunt cei care folosesc mai des date. Pentru aceasta am pus în relație utilizarea data.gov.ro, cea a altor surse de date, cea de date din propria organizație și utilizarea a măcar unui set de date cu vârsta, nivelul de educație al respondentului (facultate, master, doctorat), cu numărul de subordonați (0, sub 10, 11-30, sau peste 30) și cu dimensiunea organizației (1-10 angajați, 11-31, 31-60, 61 și peste)¹⁶.

Folosirea de date de pe data.gov.ro nu variază în funcție de vârstă și educație, dar este mai ridicată la cei ce spun că au mai mult de 30 de angajați în subordine și scade în organizațiile cu peste 30 de angajați. Educația și vârsta nu diferențiază de altfel folosirea de date pe surse ale acestora, dar cu cât educația este mai ridicată, cu atât apare o probabilitate mai mare de a declara utilizarea de date la modul general. Cu alte cuvinte, nu

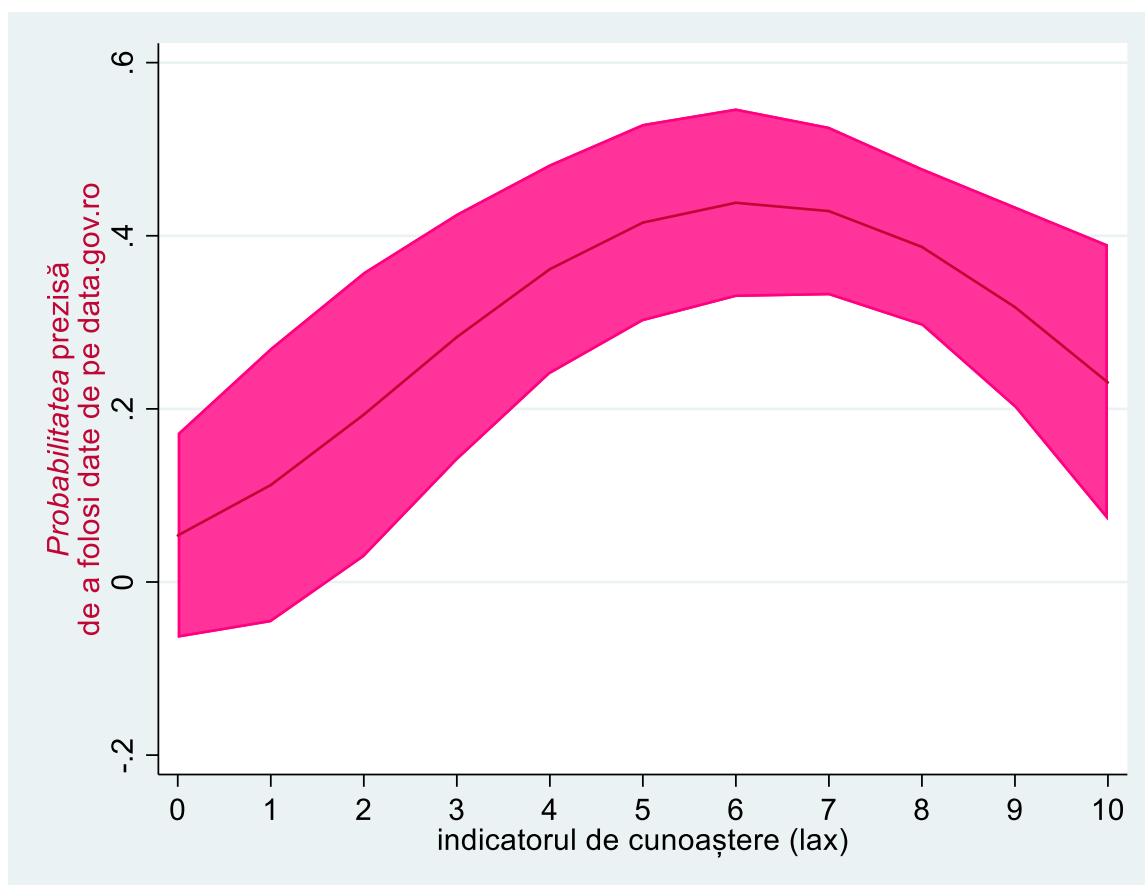
¹⁶ Am folosit în acest scop modele de regresie logistică. În ce ele ce urmează, rezultatele raportate se referă la nivelul de semnificație de .10, folosit în sensul de indicator al puterii relației, dat fiind că eșantioanele sunt în fapt nereprezentative.

contează ce educație ai pentru a utiliza orice tip de date în sine, dar contează ca predispoziție de a utiliza date. Organizațiile mai mari folosesc în mai mică măsură date produse în propria organizație sau de pe data.gov.ro. Cei care au mai mult de 30 de subalterni indică în mai mare măsură decât restul că folosesc fiecare tip de date, dar nu diferă de medie în ce privește propensiunea de a folosi date la modul general.

Paragraful de mai sus ilustrează mai degrabă confuzia în ce privește utilizarea seturilor de date. La aceasta se adaugă observația că asocierea cu gradul de cunoaștere variază oarecum neașteptat, după cum este ilustrat în Figura 22. La nivele reduse și ridicate de cunoaștere a ce înseamnă date deschise, utilizarea data.gov.ro este mai redusă. În schimb, pentru valori medii ale indicatorului de cunoaștere, probabilitatea de a utiliza seturi de date furnizate de portalul guvernamental crește spre 40%.

Același tip de relație este observat și pentru utilizarea la modul general de date. În cazul utilizării de date din alte surse sau din surse proprii, relația este similară ca formă, dar mult mai slabă (diferențele dintre indivizi în funcție de indicatorul de cunoaștere a datelor deschise sunt mult mai mici).

Figura 22. Probabilitatea de a utiliza date de pe data.gov.ro în funcție de nivelul de cunoaștere a ce înseamnă date deschise



*probabilitatea este estimată în modelele de regresie logistică menționate în text. Curba de culoare roșie reprezintă valorile estimate ale probabilității studiate, iar zona roz constituie intervalul de încredere 95%, adică plaja de valori în care s-ar situa probabilitatea estimată în 95 din 100 de cazuri dacă eșantionul ar fi reprezentativ.

MEDIUL ACADEMIC: PRODUȚIE DE DATE

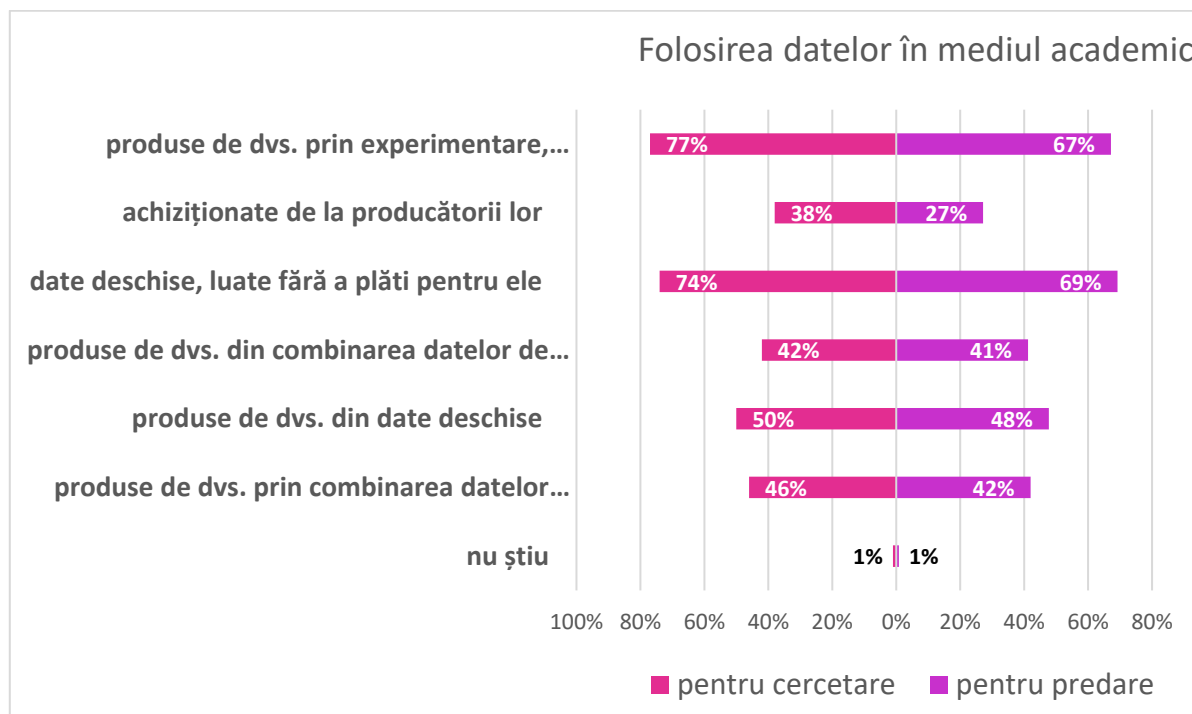
În mediul academic, utilizarea datelor este indiscutabil legată și de capacitatea de a le produce, chiar dacă nu este condiționată de aceasta. Între respondenții la sondajul cu universitari, 87% afirmă că au cel puțin o dată implicați în producere de date, 10% nu au făcut acest lucru, iar 3% nu știu ce să răspundă la întrebare. Cu alte cuvinte, între cei ce au răspuns la sondaj avem o majoritate covârșitoare ce are capacitatea de a produce date. Fără îndoială aceasta poate fi și un impediment în calea utilizării datelor deschise, pentru că este posibil ca analiza secundară de date (produse de alții) să devină mai puțin importantă. Este însă puțin probabil să se întâmple astfel, dat fiind că nu am observat o asociere pozitivă între a produce date și cunoașterea despre datele deschise.

Între cei ce produc date, jumătate (exact 50%) nu au planuri de management a datelor sau le au rar, 14% spun că au astfel de planuri în jumătate din acțiunile lor, iar 32% afirmă că au astfel de planuri des sau foarte des (4% nu au știut să răspundă la întrebare, cel mai probabil nu au astfel de planuri). De această dată nu există legătură între cunoașterea legată de date deschise și a avea planuri de management a datelor.

MEDIUL ACADEMIC: UTILIZARE ȘI REUTILIZARE

93% din eșantionul de respondenți din mediul academic declară că folosesc date. Am investigat ce fel de date folosesc, având ca țintă să aflăm dacă există o predispoziție spre a utiliza date deschise și de a le reutiliza.

Figura 23. Tipuri de date folosite în mediul academic



*Cifrele reprezintă ponderea celor ce declară că folosesc tipul respective de date.

Figura 23 ilustrează utilizarea de date în mediul academic. O primă observație este lipsa de diferențieri majore între cercetare și predare. În ambele tipuri de demersuri, predominante sunt datele produse de

cercetător în sine sau de echipele din care face parte, la care se adaugă datele deschise. Aceasta întărește potențialul de adresabilitate a portalului data.gov.ro. Recombinarea și crearea de noi de seturi de date este mai puțin întâlnită decât utilizarea directă, dar atrage și ea aproximativ jumătate din cei chestionați. Achiziția de date este mai rară, costul jucând probabil aici un rol important.

REPREZENTĂRI DESPRE CONSECINȚELE DESCHIDERII DATELOR

DESCHIDEREA DATELOR CA PROCES EVOLUTIV?

Sondajele din 2020 au inclus un pachet de itemi privind consecințele deschiderii datelor. Respondenții au fost rugați să indice în ce măsură ei cred că deschiderea ar fi conduce la (a) o trecere de utilizare de baze de date și API în administrația publică; (b) e-guvernare și digitalizare; (c) facilitează oferirea de servicii publice; (d) crește valoarea de piață a datelor; (e) facilitează standardizarea. Formularea exactă a întrebărilor este cea reprodusă în Tabelul 26.

Tabelul 26. Opinii privind consecințele deschiderii datelor: 2020

		În foarte mică măsură	În mică măsură	În mare măsură	În foarte mare măsură	Nu știu	NR	Total
Admin publică 2020	se face trecerea la utilizarea de baze de date și API-uri în administrația publică	1%	9%	27%	27%	18%	19%	100%
	constituie un pas normal pentru e-guvernare, digitalizare	2%	1%	26%	50%	5%	16%	100%
	facilitează îmbunătățirea modului în care sunt oferite serviciile publice	1%	3%	28%	48%	3%	17%	100%
	aduce valoare pe piață prin agregarea datelor din diverse surse inclusiv de la nivel european	1%	4%	27%	46%	6%	16%	100%
	facilitează standardizarea	2%	3%	26%	45%	6%	18%	100%
Academic 2020	se face trecerea la utilizarea de baze de date și API-uri în administrația publică	5%	10%	32%	20%	27%	7%	100%
	constituie un pas normal pentru e-guvernare, digitalizare	2%	4%	36%	46%	8%	4%	100%
	facilitează îmbunătățirea modului în care sunt oferite serviciile publice	2%	6%	35%	45%	8%	4%	100%
	aduce valoare pe piață prin agregarea datelor din diverse surse inclusiv de la nivel european	2%	4%	34%	46%	9%	5%	100%
	facilitează standardizarea	2%	5%	36%	42%	9%	6%	100%

Majorități confortabile identifică efecte pozitive în toate aceste privințe, cu observația unei anumite moderații la prima întrebare, unde și numărul de refuzuri de răspuns și „nu știu” a fost mult mai ridicat. Cauza este dată de prezența acronimului API în întrebare, indicând încă o dată faptul că este nevoie de comunicare care să se axeze pe un limbaj accesibil atât în ce privește administrația publică, cât și mediul academic.

Comparația dintre opiniile administrației publice și cele ale mediului academic indică un ușor optimism mai ridicat în ce privește administrația publică. Acesta poate fi și un efect al modului de selecție a eșantionului provenit din administrația publică.

Pentru a studia diferențierile în interiorul eșantioanelor, am construit un indicator unic de reprezentare pozitivă asupra consecințelor deschiderii datelor. Indicatorul include informație din toți itemii descriși în Tabelul 26, exceptându-l pe primul, dat fiind că acesta este afectat de un număr mare de non-răspunsuri¹⁷. Analiza¹⁸ nu relevă diferențieri esențiale în interiorul eșantioanelor, cu două excepții. Pe de o parte, cunoașterea despre date deschise acționează drept catalizator pentru optimismul în privința consecințelor, până la un punct acesta crescând pe măsură ce crește cunoașterea despre date deschise. Dar cei care știu mai multe despre date deschise devin mai circumspecți în privința consecințelor. Pe de altă parte, sunt două domenii ale științei (medicina și ingineria) în care respondenții sunt mult mai reținuți în a fi optimiști în ce privește consecințele deschiderii datelor, în timp ce la universitarii din domeniu dreptului optimismul este mai puternic. Consecința este legată de nevoia de a orienta campanii mai atente de comunicare către cei din medicină și drept.

MEDIUL ACADEMIC: CONSECINȚE ALE DATELOR DESCHISE ASUPRA INTEGRITĂȚII ACADEMICE

Deschiderea datelor este asociată ca principiu general transparenței. În chestionarul pentru mediul academic am propus două baterii distincte de itemi care discută despre transparență și consecințe ale sale în ce privește integritatea academică.

Prima baterie se referă la consecințe explicite ale deschiderii datelor. Respondenții consideră în majoritate că deschiderea datelor previne fraudă academică (79%), permite încurajarea integrității academice (79%), și că ajută studenții să evite plagiatul (70%). Din aceste întrebări am putut realiza un indicator compozit ce reflectă nivelul în care respondentul crede că deschiderea datelor facilitează integritatea academică¹⁹. Analizând diferențele între diferite categorii de respondenți²⁰, am constatat că doar gradul didactic aduce modificări în ce privesc beneficiile așteptate de la deschiderea datelor: cu cât gradul didactic crește, de la asistent la lector, apoi la conferențiar și în final la profesor, cu atât reprezentarea că datele deschise au consecințe pozitive asupra integrității academice este mai puternică. Este de observat că în această privință, cunoașterea despre date deschise nu are niciun efect.

A doua baterie este una de integritate academică, constând în respingerea plagiatului, a prezentării de date falsificate, a raporta ceea ce cere clientul nu ceea ce observi, a nu prezenta rezultatele ce resping propria teorie, a prezenta ca certe idei pentru care nu ai dovezi. Și în acest caz, am construit un factor de integritate academică²¹. Am studiat²² asocierea dintre familiaritatea cu datele deschise (indicatorul de cunoaștere) și integritatea academică. Nici din acest punct de vedere nu am constatat vreo consecință a datelor deschise.

¹⁷ Indicatorul este construit ca scor factorial. Analiza factorială (extracție prin maximum likelihood) explică 73% din variația totală. Comunalitățile sunt peste 0.67. KMO=.840. Testul grohotișului (scree plot) indică prezența unui singur factor.

¹⁸ Modele de regresie (OLS), controlând pentru variabilele folosite și în celelalte capitole.

¹⁹ Indicatorul este calculat ca scor factorial (KMO=0.718, toate comunalitățile sunt mai mari de 0.5, iar variația explicată este de 72%, factorul fiind extras folosind principal Axis Factoring. Scalele inițiale erau de patru puncte.

²⁰ Prin analiză de regresie.

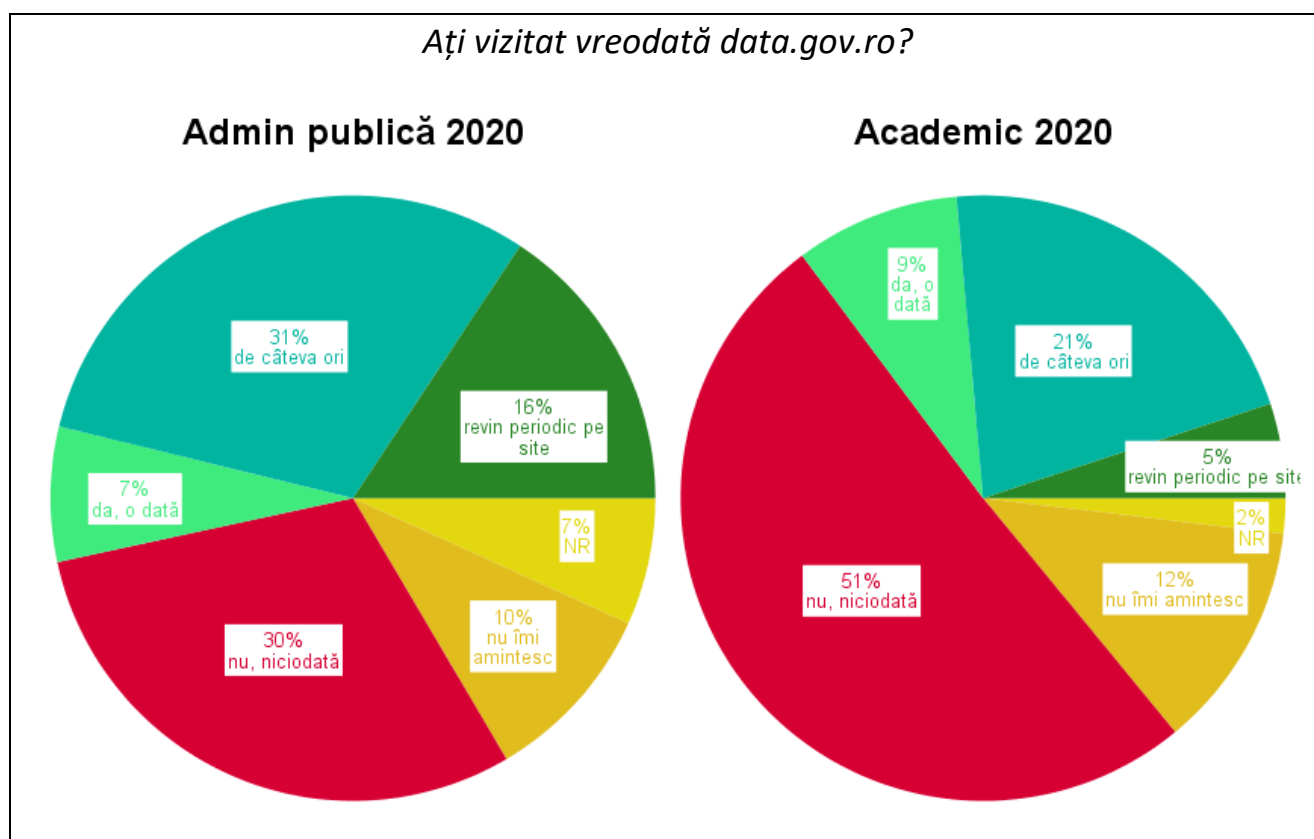
²¹ Analiză factorială, Maximum Likelihood, comunalități mai mari de 0.1, 8 itemi scalați de la 1 la 8, testul grohotișului (scree plot) semnalează prezența unui singur factor, KMO=0.737, factorul extras explică 29% din variația totală.

²² În modele de regresie.

DATA.GOV.RO

Utilizarea data.gov.ro este redusă în rândul celor ce au răspuns la sondaje. Figura 24 relevă că sunt puțini cei care accesează data.gov.ro: aproximativ jumătate (54%) dintre cei ce au răspuns la chestionarul adresat administrației publice și o treime dintre cei din mediul academic (35%).

Figura 24. Utilizarea data.gov.ro de către respondenții l sondajele din 2020



Care dintre următoarele se aplică vizitelor dvs. pe data.gov.ro

Adm pub 2020		Academic 2020
17%	[Descărcarea de date]	13%
23%	[Curiozitatea]	23%
16%	[Verific dacă sunt actualizări ale unor date descărcate]	9%
4%	[Căutare de meta-data pentru date deja descărcate]	4%
15%	[Upload de date]	0%
6%	[Alt motiv]	2%

* Cifrele din tabelul cu motivele utilizării se referă exclusive la cei ce au vizitat platforma și nu sunt însumabile: fiecare respondent a putut selecta mai multe motive de a vizita site-ul.

Motivele de accesare a site-ului sunt destul de rar legate de utilizarea datelor (Figura 24). În plus față de cifrele din figură, cei din mediul academic au fost întrebați explicit dacă au folosit vreodată date preluate de pe data.gov.ro. 14% au indicat faptul că au făcut acest lucru. Pe domenii, ei provin din practic toate ramurile științei.

Dincolo de utilizarea efectivă, respondenții au fost întrebați ce seturi de date noi ar dori să fie publicate pe data.gov.ro. A rezultat o listă lungă de câteva pagini, pe care o reproducem ca atare în Anexe (pagina 111). Varietatea prezentă în listă este remarcabilă, subliniind nevoia de date existentă. Dincolo de ea, este de observat faptul că o parte din cererile de date sunt deja satisfăcute. De aici observația despre necesitatea de a populariza platforma, mai ales în rândul mediului academic, pentru a-i crește valoarea de utilizare, aducând datele mai aproape de utilizatorii potențiali.

Figura 25. Publicarea de date pe data.gov.ro

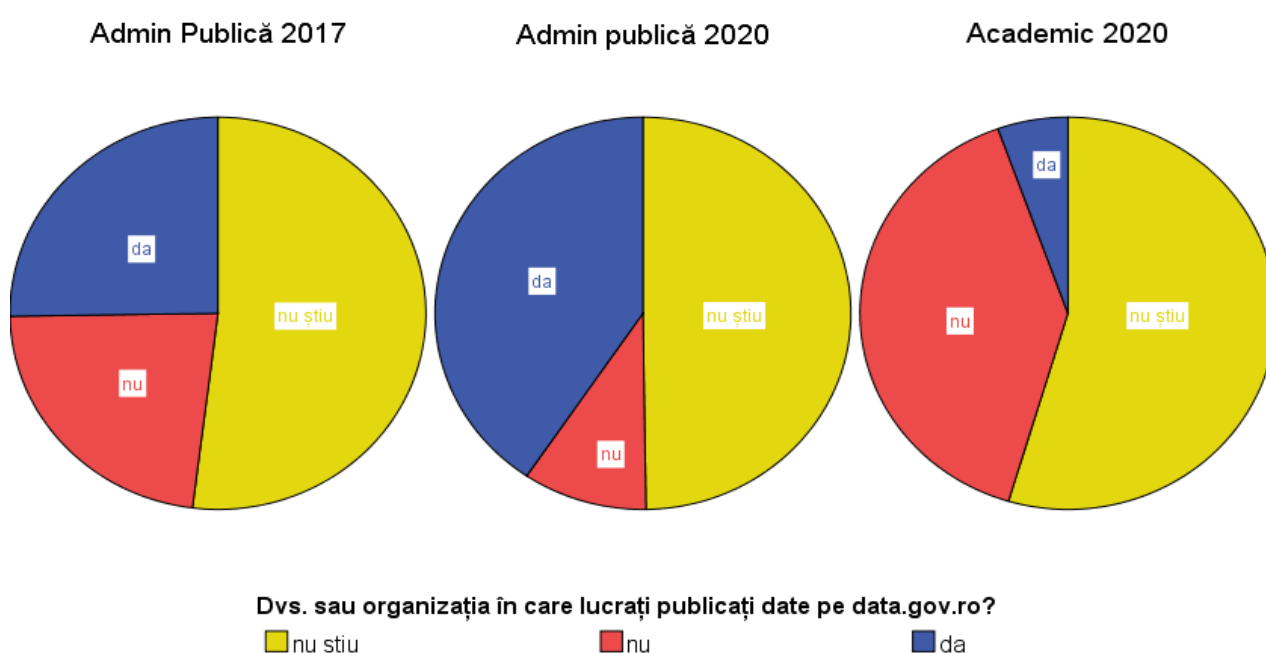


Figura 25 pune în evidență diferența notabilă dintre eșantioane cu privire la publicarea de date pe data.gov.ro. Dacă pentru eșantionul de respondenți din mediul academic aceasta este o practică rară, în cazul administrației publice constatăm o creștere substanțială între 2017 și 2020²³.

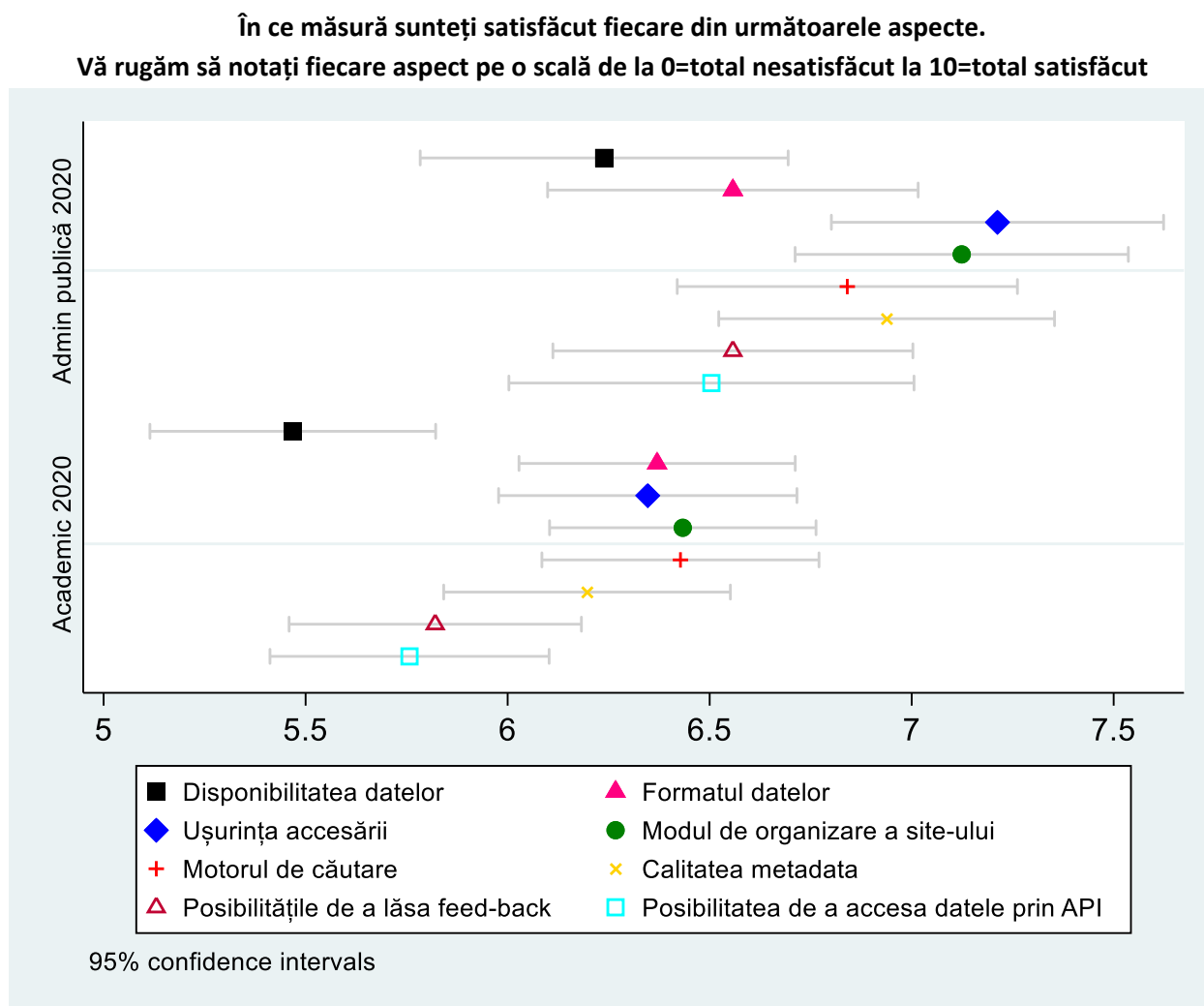
Totuși, numărul respondenților care nu sunt informați cu privire la comportamentul organizației din care fac parte continuă să fie aproximativ jumătate din total. Rezultă nevoia de a discuta în interiorul agențiilor administrației publice, indiferent de natura și localizarea lor (centrale, regionale, județene, locale etc.), despre fluxurile de date și posibilitatea de a aduce în spațiul public datele produse în interiorul administrației publice.

²³ Regresia multinomială confirmă rezultatele univariate.

REZULTATE DIN SONDAJE

Sondajele din 2020 au inclus, pentru cei ce au accesat cel puțin o dată data.gov.ro, o baterie de itemi ce a vizat evaluarea platformei. Cele opt aspecte avute în vedere sunt prezentate grafic în Figura 26.

Figura 26. Satisfacția cu diferite elemente ale data.gov.ro



*Cifrele reprezentate grafic prin markerii colorați sunt medii pentru cele două eșantioane. Cu titlu informativ sunt prezentate și intervalele de încredere ale estimatorilor respectivi. Când acestea se suprapun, dacă eșantioanele folosite ar fi probabilistice, am spune că nu sunt diferențe semnificative.

Rezultatele relevă o satisfacție peste mijlocul scalei în ce privește fiecare dintre aspectele luate în considerare. Ușurința accesării, modul de organizare a site-ului, motorul de căutare sunt elementele ce ies în evidență atât în ce privește reprezentările provenite din administrația publică, cât și pe cele din mediul academic. La polul opus, pentru ambele categorii se află disponibilitatea datelor. Impresia de ansamblu rămâne că avem de a face cu satisfacție mai ridicată cu forma site-ului și cu opinii mai critice în ce privește conținutul. Din acest punct de vedere data.gov.ro poate fi satisfăcut cu starea curentă: comunități critice dar apreciative la adresa efortului constituie un pas uriaș înainte. Implicația este legată de necesitatea de a atrage seturi de date suplimentare pentru perioada imediat următoare. Acestea este consonant și cu multitudinea

răspunsurilor despre noi seturi de date ce sunt solicitate, după cum relevă a doua secțiune din Anexe („Ce seturi noi de date sunt solicitate pe data.gov.ro”, începând de la pagina 111 a acestui volum).

Comparând cele două eșantioane, este evidentă discrepanța aprecierilor: administrația publică este mult mai decât mediul academic. Practic, fiecare dintre cele opt aspecte puse în discuție înregistrează un scor mediu de satisfacție în eșantionul administrației publice cu un punct mai mare decât în eșantionul provenit din mediul academic. Diferența de un punct este mare deja dacă o raportăm la scala de 11 puncte (având dimensiunea de $10-0=10$ puncte) care a fost folosită în chestionare. Raportat la mediile observate, acest punct diferență este încă și mai mare.

Ca și în alte analize din acest volum, apare nevoia de a focaliza pe viitor atenția legată de promovarea data.gov.ro și în afara administrației publice.

REZULTATE DIN CHESTIONARUL DE SATISFAȚIE IMPLEMENTAT PE SITE

Chestionarul de satisfacție a fost recomandat în fazele inițiale ale cercetării de față și este descris în capitolul *Chestionare de feedback asociat data.gov.ro* al acestui volum (începând cu pagina 100). El a fost implementat în platforma data.gov.ro în decembrie 2020. În perioada 20 decembrie 2020 – 4 martie 2021, conform Google Analytics (a se vedea capitolul *Back-log analysis*, începând cu pagina 22 a acestui volum), site-ul data.gov.ro a fost accesat de 28.885 utilizatori. Dintre aceștia, 1024 au accesat chestionarul, care apare ca pop-up pe site. 305 au completat întreg chestionarul de satisfacție, 470 nu au completat nimic, iar restul de 250 au completat parțial chestionarul.

Figura 27. Număr de completări pe zile ale chestionarului de satisfacție disponibil pe data.gov.ro în perioada 20 decembrie 2020-4 martie 2021

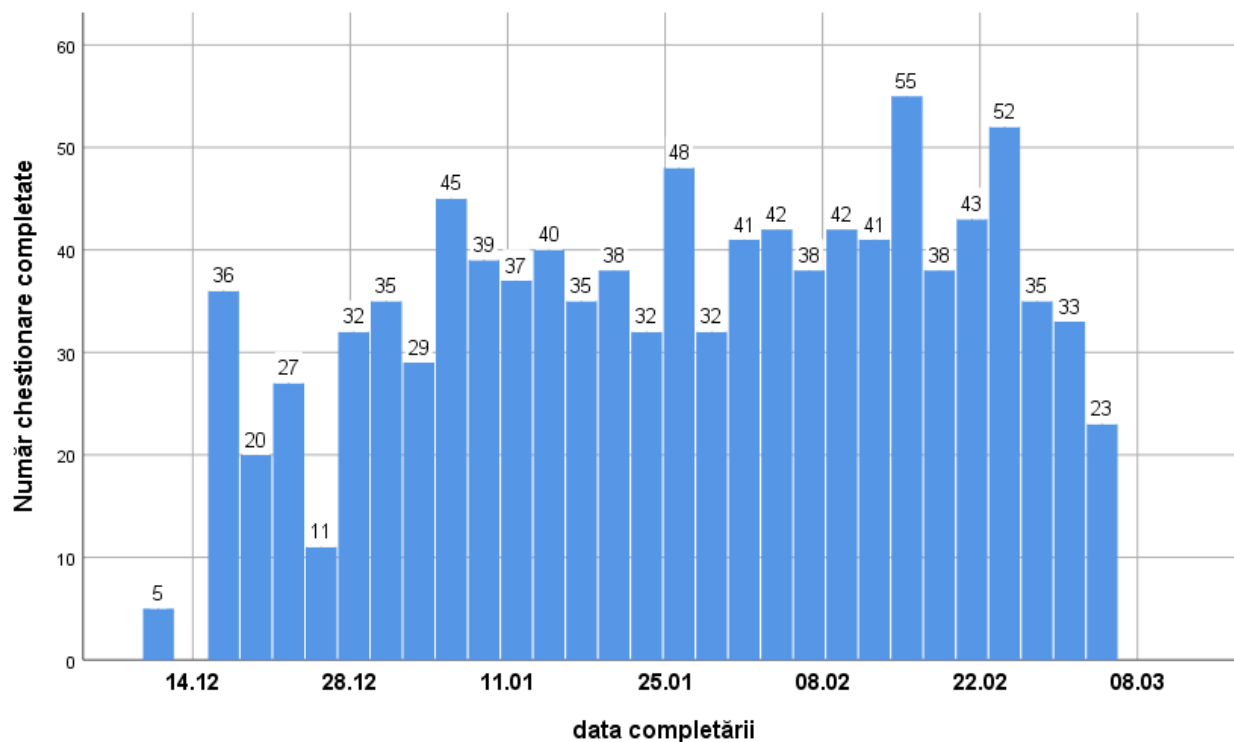
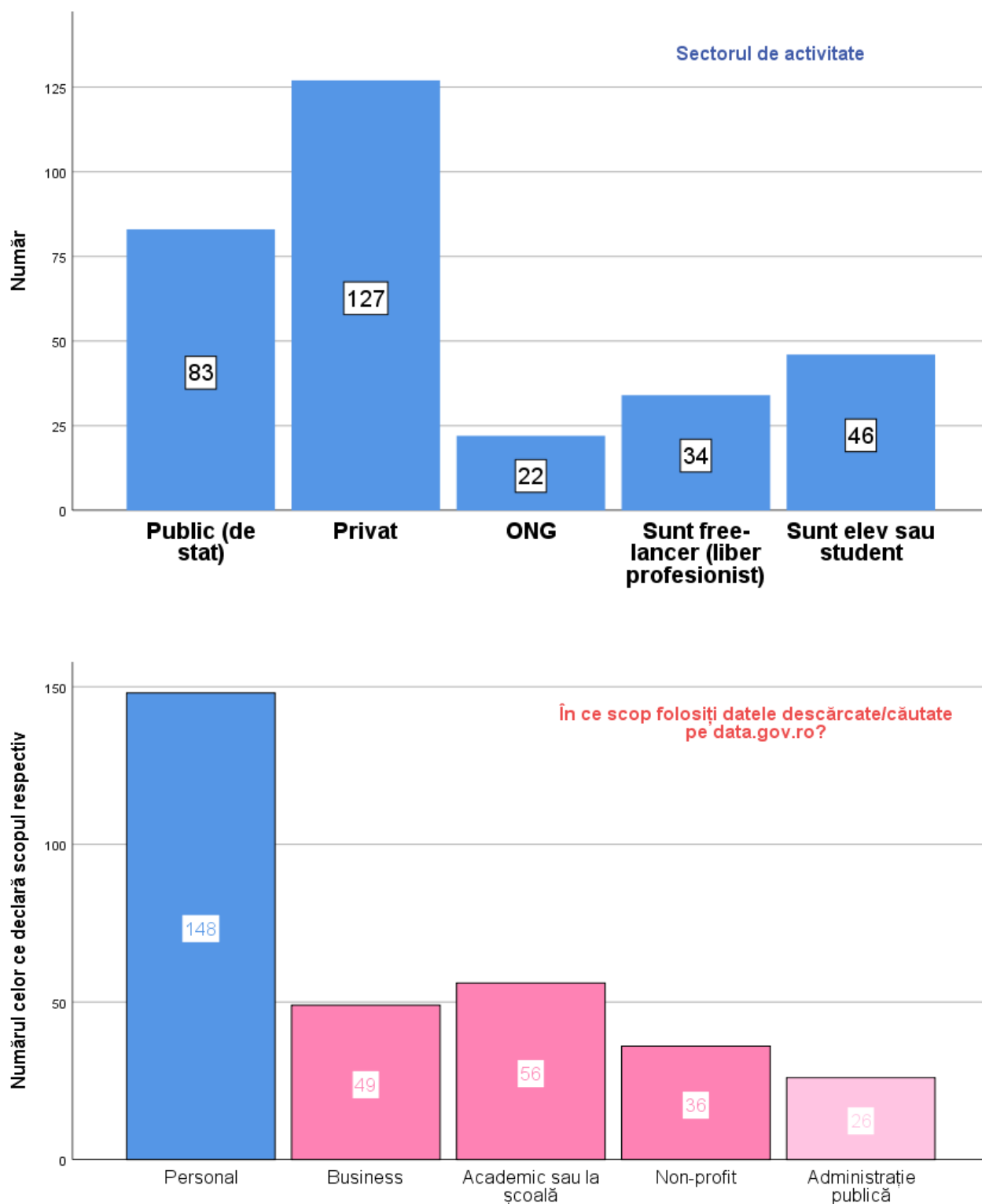


Figura 28. Detalii privind cei ce au completat chestionarul de satisfacție



Întrebați despre intențiile cu care au ajuns pe data.gov.ro, 2% dintre cei ce au răspuns la întrebare (în total 642) au indicat descărcarea de date, 26% curiozitatea, 11% căutarea de actualizări, 3% căutarea de metadate pentru date deja descărcate, iar 21% au spus că au accesat data.gov.ro din alt motiv.

Gradul de satisfacție cu diverse aspecte ale site-ului este relativ similar satisfacției observate pentru respondenții la sondajul realizat cu mediul academic.

Tabelul 27 Satisfacția cu data.gov.ro

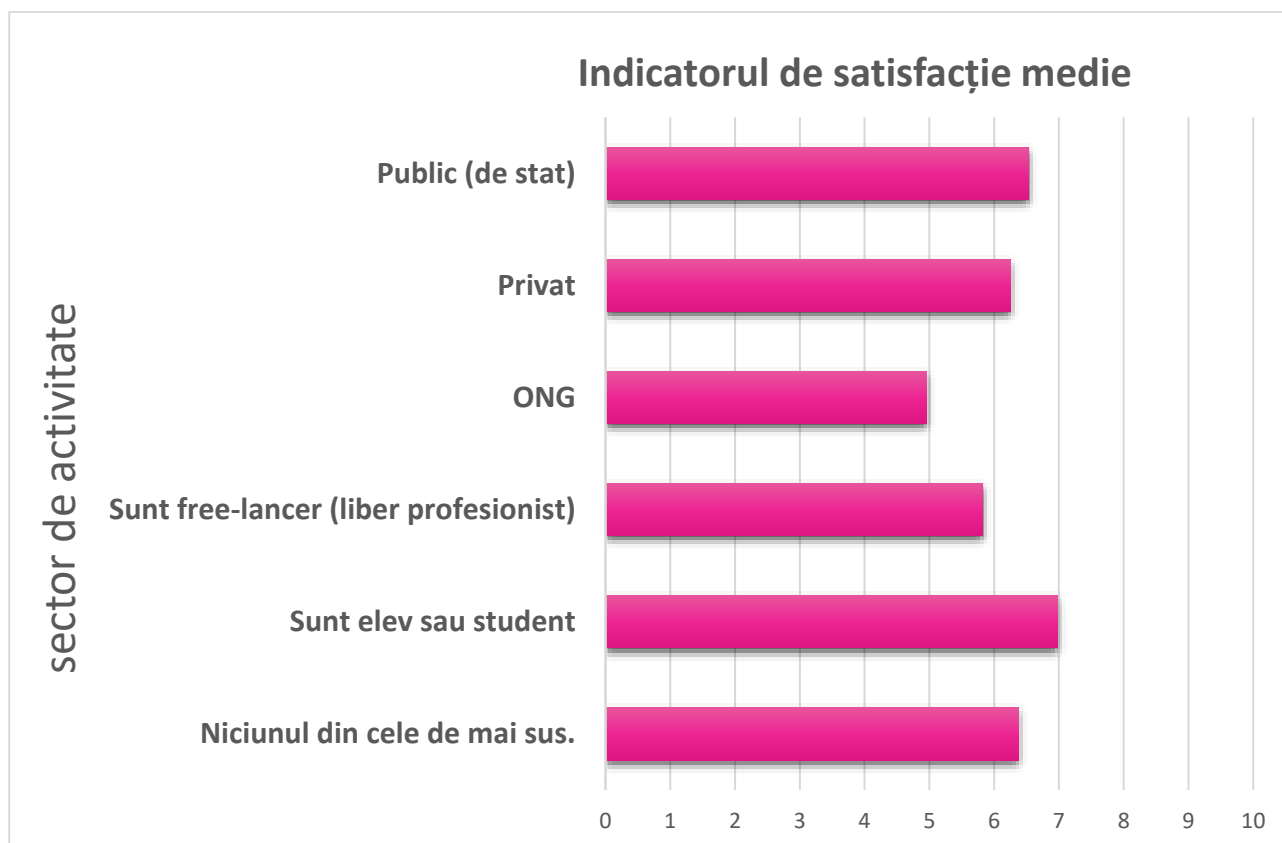
În ce măsură sunteți satisfăcut de fiecare din următoarele aspecte. Vă rugăm să notați fiecare aspect pe o scală de la 0=total nesatisfăcut la 10=total satisfăcut, putând alege orice valoare intermediară.

	N*	Minim	Maxim	Media	Abaterea standard
SAT_1 Disponibilitatea datelor	300	0	10	6,04	3,134
SAT_2 Formatul datelor	290	0	10	6,26	3,122
SAT_3 Ușurința accesării	293	0	10	6,54	3,091
SAT_4 Modul de organizare a site-ului	292	0	10	6,23	3,092
SAT_5 Motorul de căutare	289	0	10	6,36	2,986
SAT_6 Calitatea metadata	288	0	10	6,18	3,019
SAT_7 Posibilitățile de a lăsa feed-back pentru un set de date	289	0	10	6,23	3,072

*N este numărul de răspunsuri obținute la întrebarea respectivă.

Pentru fiecare respondent am calculat o medie a indicatorilor din Tabelul 27²⁴. Figura 29 prezintă rezultatele în funcție de sectorul de activitate al respondentului. Se observă că diferențele sunt mici, dar free-lancer-ii și mai ales cei din sectorul ONG prezintă nivele ale satisfacției mai reduse.

Figura 29. Satisfacția utilizatorilor data.gov.ro în funcție de sectorul de activitate

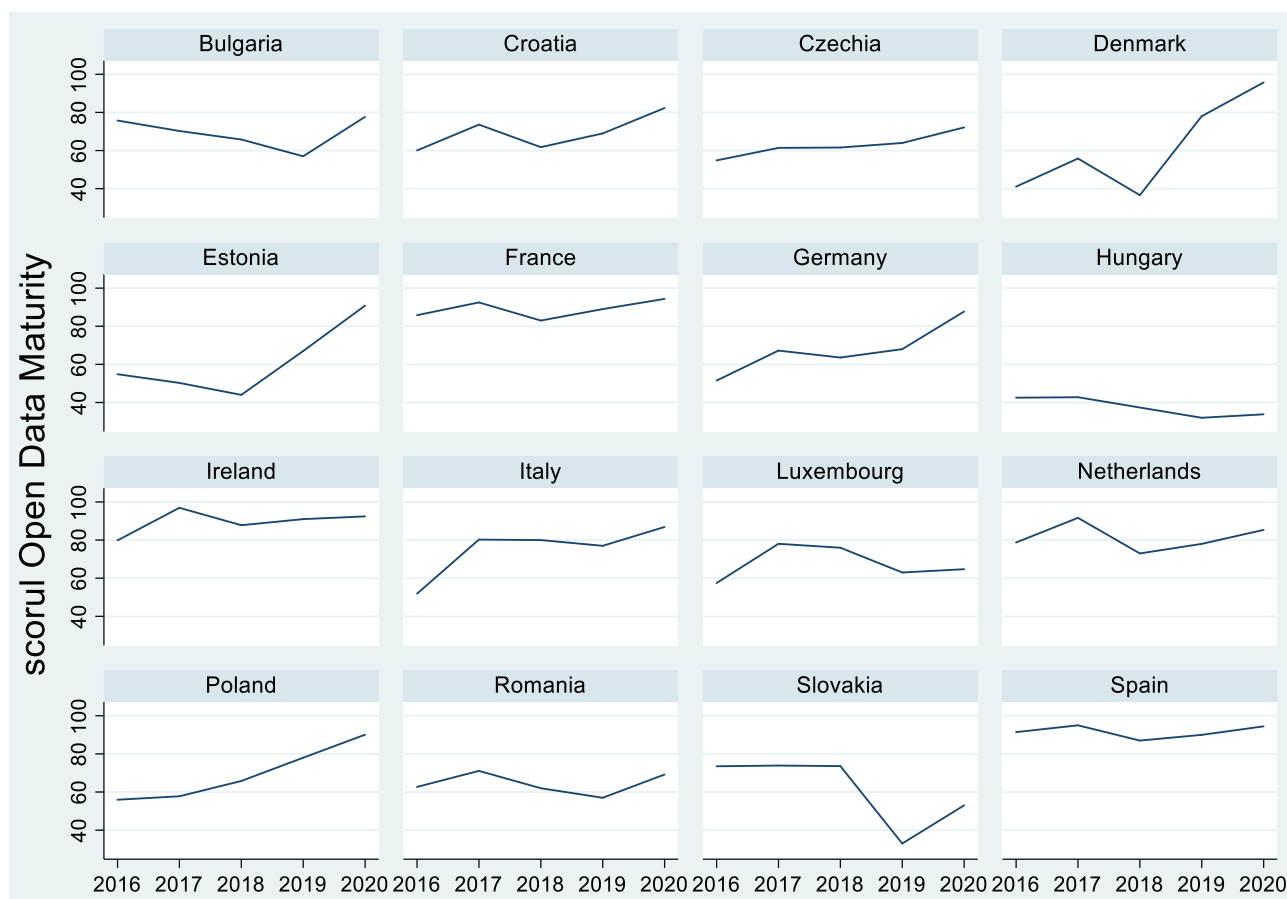


²⁴ Analiza factorială confirmă faptul că indicatorul este legitim a fi calculat.

ANALIZA COMPARATIVĂ A IMPACTULUI

Un mod de a evalua impactul deschiderii datelor este de a observa asocierea dintre dezvoltarea economică și gradul de deschidere a datelor. Am folosit pentru aceasta indexul Open Maturity Data, produs anual de către European Data Portal²⁵. Figura 30 prezintă evoluția indexului în câteva țări europene. Am selectat pentru grafic țările vecine României, țările mai mari din Europa și performerii în acest top pentru 2020.

Figura 30. Evoluția Open Maturity Data între 2016 și 2020 în câteva țări europene



*Scorul este calculat ca pondere din maximumul posibil, cu observația că metodologia de calcul suferă modificări în 2018 și 2020.

Am pus acest indicator în relație cu rata de creștere a PIB în anul imediat următor. Am realiza o analiză panel, în care am luat în calcul și nivelul PIB pe locuitor. Am repetat apoi analiza ținând cont și de abilitățile digitale ale populației țărilor europene și ponderea cheltuielilor publice în PIB²⁶.

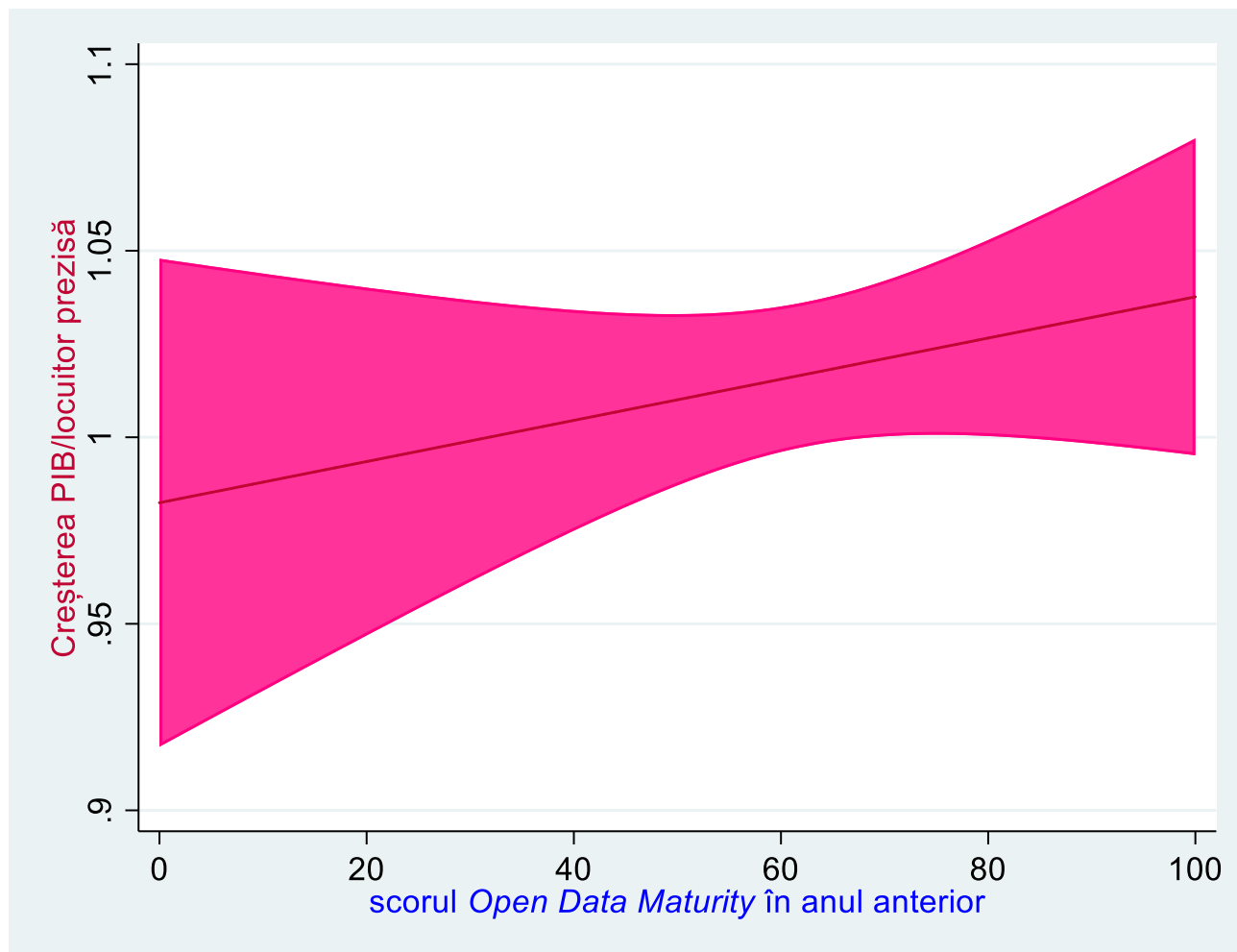
Rezultatul este simplu: creșteri ale scorului de maturitate a datelor, relativ la maximumul posibil al acestuia, conduc la creșteri mici ale GDP pe capita. Mai exact, o creștere de un punct procentual a Open Maturity Data se reflectă în medie într-o creștere a PIB/locuitor cu 0.0005 procente. Creșterea este stabilă și dacă luăm în calcul și ceilalți indicatori menționați și este reprezentată grafic în Figura 31. Aparent discutăm de o modificare minusculă a GDP/locuitor. Să remarcăm însă că există numeroși factori care determină PIB/locuitor și orice creștere mică este importantă. Mai mult, este vorba de o creștere bazată pe o alocare

²⁵ <https://www.europeandataportal.eu/en/dashboard/2016>.

²⁶ E vorba de analize de date panel (fixed effects models).

bugetară minusculă, așa cum este cea dedicată deschiderii datelor. Concluzia este simplă: deschiderea datelor împinge înainte economia.

Figura 31. Impactul creșterii scorului Open Data Maturity asupra creșterii PIB/locuitor



*Rezultate bazate pe analiza datelor panel. Pentru detalii asupra modului de calcul, a se vedea textul acestui capitol.

Concluzia se aplică mediei țărilor europene. Cum însă România este parte a acestei familii de țări, și în fapt acest tip de analiză compară țările între ele, se poate aplica concluzia de mai sus și României. Implicația este dată de nevoia de a dezvolta deschiderea datelor, ca pe un pariu care stimulează creșterea economică.

CONCLUZII ASUPRA IMPACTULUI

Pe ansamblu, avem de a face cu o creștere a notorietății și cunoașterii datelor, dar aceasta este limitată la nivelul administrației publice. O strategie de creștere a impactului la nivelul altor populații țintă poate constitui pasul următor în politicile de deschidere a datelor și de promovare a data.gov.ro, fiind element cheie în a crește efectele de multiplicare și în a încuraja utilizarea și reutilizarea datelor.

Legitimitatea este ridicată, fiind însă în strânsă relație cu cunoașterea. Legitimitatea este deopotrivă a deschiderii datelor ca principiu și este acordată și data.gov.ro. În cazul data.gov.ro, legitimitatea derivă din caracterul integrator al platformei.

Deschiderea spre a deschide date și spre a le publica pe data.gov.ro este și ea în creștere.

Utilizarea este restrânsă pentru moment, depinzând de capacitatea societății de a utiliza date, care la rândul ei condiționează și disponibilitatea datelor. Un efort de a oferi modele se impune în acest sens. El presupune o investiție minimală, iar efectele de multiplicare pot fi extrem de puternice.

Tabelul 28 sintetizează principalele rezultate într-un mod ușor de citit. Documentarea fiecărei concluzii a fost realizată în capitolele anterioare. Sunt câteva domenii mari în care se poate observa un impact al data.gov.ro și a deschiderii datelor în România:

- Creștere a notorietății și cunoașterii
- Creștere a legitimității deschiderii datelor
- Creștere a deschiderii spre disponibilitate nelimitată (gratuită)
- Dorință a operatorilor din afara domeniului public de a (re)utiliza datele, adesea transpusă în comportamente de (re)utilizare, inclusiv în scop comercial
- O creștere notabilă a celor care acceptă ca pregătirea pentru publicare a datelor vechi să fie realizată de către terți
- Creștere a timpului petrecut pe data.gov.ro și a numărului de pagini accesate într-o sesiune unică
- Evaluare apreciativă a data.gov.ro, mai ales din partea administrației publice, dar și o atitudine critică față de disponibilitatea seturilor de date

Analiza realizată permite identificarea câtorva nevoi certe de intervenție:

- Notorietatea data.gov.ro în mediul academic
- Potențial de preluare a datelor din mediul academic
- Grupa de vârstă mai tânără (mai ales sub 30 de ani) manifestă o nevoie de formare în ce privește elemente precum: notorietate, deschidere spre deschidere necondiționată
- Organizații medii (10-60 de angajați): sunt ușor mai reticente decât restul la a furniza date în mod gratuit/universal.
- nevoia de profesionalizare a datelor deschise, mai ales sub aspectul creșterii calității, dar și a capacității de utilizare în interiorul administrației publice
- cunoașterea acționează ca element catalizator al legitimității, deschiderii spre deschiderea datelor, utilizarea datelor etc.

- este o nevoie de a oferi modele de deschidere a datelor, centrate pe fluxurile practice de deschidere la nivel de organizație și specificarea clară a persoanelor ce o pot face. În acest sens este legitimat puternic rolul unui organism central, în speță al SGG.
- Utilizarea redusă a datelor în mediul administrației publice constituie probabil o barieră importantă în calea deschiderii datelor: pe de o parte nu creează cerere, pe de alta nu conduce la înțelegerea procesului de deschidere a datelor.
- În mediul academic, ancore ale deschiderii datelor pot fi profesorii universitari mai tineri (cei cu gradul de profesor). Nevoia de formare este mai acută la asistenți universitari.
- Nevoia de formare este reclamată de trei sferturi din cei ce au răspuns la chestionare, referindu-se deopotrivă la utilizarea și la publicarea datelor

Tabelul 28. Tabel sinoptic al principalelor rezultate

Notorietate	La nivelul administrației publice, cunoașterea despre date deschise este substanțial mai ridicată în 2020 decât în 2017, iar informația vine dintr-un portofoliu variat de surse, atestând circulația acesteia. În schimb, cunoașterea la nivelul mediului academic este marcată de multe ambiguități.
	Sunt în continuare dificultăți de a înțelege unii termeni (exemplu: API). O strategie de comunicare axată pe limbaj natural este recomandabilă.
	Reprezentarea celor intervievați în sondaje asupra gradului de publicare a datelor deschise în România prin raportare la restul Europei este majoritar pesimistă. Optimiștii sunt puțini, iar cei care nu au informație despre alte țări reprezintă de asemenea o pondere mare. Pesimismul este mai pregnant în mediul academic. Reprezentările respective vin în contradicție cu rapoartele comparative internaționale care poziționează România în media europeană. Apare astfel o nevoie certă pentru comunicare mai bună și promovare a datelor deschise și a data.gov.ro.
	Experții în date deschise consultați sunt și ei de părere că notorietatea este sub potențial și notează lipsa de implicare a guvernului în ansamblul său în a legitima data.gov.ro.
	Studiile de caz relevă și ele că informația despre data.gov.ro ajunge uneori greu chiar și la cei care caută date deschise.
	Nivelul cunoașterii datelor deschise este asociat cu informarea despre data.gov.ro. Cele două acționează probabil ca și catalizatoare una pentru cealaltă.
Legitimitate	Experții participanți la panelul Delphi asociază legitimitatea cu caracterul integrator al data.gov.ro și cu actualizarea datelor. Discuția s-a construit mai ales în jurul relației cu alte portaluri și cu nevoia de a te descurca cu surse multiple de date.
	În administrația publică apare un efect de creștere a acceptării reutilizării datelor deschise, iar creșterea este mai puternică la cei ce știu mai puțin despre datele deschise. Aceste efect în sine constituie un semn remarcabil al schimbării la nivel societal și creează premisele sustenabilității deschiderii, utilizării și reutilizării datelor.
	În mediul academic apare o susținere medie a difuzării gratuite a datelor deschise produse, dar respondenții alocă sarcina difuzării datelor în mică măsură propriei organizații. Rezultă o posibilă nișă pentru data.gov.ro de a prelua date din mediul academic și a le reda circulației în spațiul public, dar și o nevoie de a crește legitimitatea difuzării datelor și pentru utilizarea lor comercială.

	<p>Cunoașterea despre date deschise este asociată unei legitimități mai accentuate a deschiderii acestora.</p> <p>În mediul academic, vectorii esențiali spre deschidere se dovedesc a fi profesorii universitari (ca grad didactic) mai tineri. Apare o nevoie explicită de a crește legitimitatea deschiderii datelor în Arte, Științe Umaniste și Drept, dar acestea sunt tocmai domeniile care produc (și folosesc) în mai mică măsură date.</p>
Disponibilitate	<p>Reprezentările administrației publice pun în 2020 mai explicit accent pe elemente intrinseci deschiderii datelor ca factori de catalizare a disponibilității acestora. Este un indiciu asupra unei noi etape în procesul de deschidere a datelor, în care încep să primeze preocupări pentru calitate și capacitate de deschidere și mai puțin pe voința politică de a deschide datele. Aceasta este simultană cu o recunoaștere a expunerii publice mai ridicate și a eforturilor SGG în acest sens.</p> <p>Deschiderea spre a face disponibile datele este mai ridicată decât în 2017. Nu am observat diferențe între administrația publică și mediul academic. Deschiderea spre disponibilitate este mai mare la cei ce știu mai multe despre date deschise. Dorința de a monetiza datele apare mai ales la generația sub 40 de ani, este mai puternică la cei sub 30 de ani, și este mai redusă la cei între 40 și 49 de ani. La toate grupele de vârstă există însă majorități ce sunt în favoarea deschiderii nelimitate a datelor.</p> <p>Una dintre chestiunile practice cu care se confruntă producătorii de date este cine să realizeze deschiderea și publicarea acestora. La acest nivel persistă confuzia între rolul departamentelor de relații publice și comunicare și a celor de IT, precum și în ce privește rolul personal în procesul de deschidere. Pe această dimensiune nu se observă schimbări substanțiale față de 2017.</p> <p>Aceeași necunoaștere apare în distribuirea rolurilor și înțelegerea sarcinilor organizațiilor din cadrul administrației publice, existând un recurs majoritar mai puternic decât în 2017 la o abordare top-down.</p> <p>Deși există preocupări mai degrabă reduse de a aduce deschide datele mai vechi aducându-le în format, se observă o creștere remarcabilă a ponderii celor care văd în utilizarea terților o soluție pentru pregătirea acestor date ca să fie readuse în spațiul public.</p> <p>Există o nevoie explicită de formare în deschiderea datelor deschise care se reflectă deopotrivă în publicare și în utilizare și este prezentă și la cei din administrația publică și la cei din mediul academic. Formarea devine un element cheie în acest moment al deschiderii datelor.</p> <p>Open Data Directive este practic necunoscută administrației publice. Aceasta din urmă evaluează mai degrabă sceptic posibilitatea de a-și asuma obligații suplimentare în ce privește publicarea datelor.</p>
Utilizare	<p>În cadrul funcționarilor publici, utilizarea efectivă a datelor, indiferent de sursa acestora din urmă, este mai degrabă redusă.</p> <p>În mediul academic, unde majoritatea este și producător de date, ponderea celor care invocă planuri de management a datelor este redusă.</p> <p>Utilizarea data.gov.ro este în creștere din punct de vedere al timpului petrecut pe site și a numărului de pagini accesate. În absența unor metrici privind numărul de descărcări este însă dificil de evaluat evoluția reală. Site-ul are un design destul de neprietenos, dar include și elemente care îl fac superior mediei europene. Pe ansamblu, este comparabil</p>

	cu media platformelor similare din UE, ceea ce lasă spațiu de îmbunătățire, dacă sunt alocate resurse în acest sens.
	Publicarea pe data.gov.ro devine o practică mai frecventă, cel puțin în reprezentările celor din administrația publică
	Utilizarea din afara sectorului public este dependentă de cunoaștere. Studiile de caz realizate relevă însă deopotrivă interes, dispoziție spre reutilizare, identifică practici de reutilizare și un potențial pentru shareability.
	Utilizarea de date în mediul academic este extrem de frecventă, fiind preferate datele proprii, dar la egală importanță cu cele deschise. Aceasta creează un potențial interesant pentru data.gov.ro. Recombinarea datelor în noi seturi de date este de asemenea destul de întâlnită, deși mai rară decât utilizarea primară și secundară a datelor.
	Satisfacția cu forma și conținutul data.gov.ro este destul de ridicată și între funcționarii publici, dar și în mediul academic. Prima categorie menționată este în general mai satisfăcută decât mediul academic, dar acesta poate fi și un efect al modului de selecție a eșantionului. Se observă în plus o atitudine mai critică față de disponibilitatea datelor și una mai apreciativă față de formatul site-ului.
	Deschiderea datelor conduce la creștere economică, după cum o indică analiza comparativă a datelor pentru toate țările europene.

DIRECȚII DE ACȚIUNE

Direcțiile acțiune propuse sunt sintetizate pe șase dimensiuni distincte :

- Utilizarea nișelor existente
- Creșterea gradului de cunoaștere
- Creșterea standardelor și calității datelor publicate
- Creșterea numărului de seturi de date
- Promovarea reutilizării – exemplu de bune practici, stimulent....
- Oferirea de modele de utilizare în interiorul organizației

Aceste direcții sunt concretizate în recomandările de programe din secțiunea următoare. În final, discutăm despre instrumentele de monitorizare ce pot fi folosite pe viitor.

RECOMANDĂRI DE PROGRAME

DEZVOLTAREA PLATFORMEI ȘI CREȘTEREA CALITĂȚII DATELOR

Așteptarea cea mai importantă a utilizatorilor rămân în continuare atingerea obiectivului principal al existenței platformei data.gov.ro: publicarea de seturi de date cu calitate și valoare ridicată. Ambele dimensiuni au nevoie de acțiuni hotărâte, cât mai repede.

În ce privește funcționalitatea data.gov.ro, utilizatorii au semnalat reticența în a folosi două instrumente importante puse la dispoziție de platformă: API pentru accesarea datelor și mecanismul de feed-back pentru a contribui la îmbunătățirea conținutului. Această reticență este bazată fie pe neîncredere în buna funcționare a instrumentelor, fie pe situații concrete în care au încercat să le utilizeze și au întâmpinat probleme. Cel mai probabil, un upgrade al platformei este necesar. Cu siguranță, este nevoie de discuții aplicate cu cei care deja folosesc date, pentru a vedea exact care sunt punctele slabe și ce trebuie îmbunătățit.

În ce privește calitatea datelor, toate instrumentele folosite confirmă ipotezele cunoscute: este nevoie de mai multe date cu valoare adăugată mare; seturile de date să fie complete și actualizate la timp. Este simptomatic că există utilizatori care încep prin a folosi data.gov.ro și apoi preferă să găsească diverse metode de a prelua seturi de date de la instituțiile publice furnizoare. Acest lucru înseamnă că data.gov.ro reușește să contribuie la notorietatea datelor deschise, dar nu reușește întotdeauna să ofere un serviciu stabil pe care să poate fi construit un proiect de reutilizare.

DATA.GOV.RO CA ELEMENT INTEGRATOR

Prezența unor platforme ale administrației publice dedicate pentru tipuri de date specifice unui domeniu îngust, independent de data.gov.ro, poate fi privită drept consecință a existenței platformei guvernamentale și ca provocare la adresa acesteia. Pe de o parte, este vorba de un impact autentic, data.gov.ro fiind probabil exemplul de urmat în domeniu. Pe de alta, înseamnă o nevoie a utilizatorilor de a căuta date pe platforme multiple.

Pentru a-și păstra legitimitatea și a crește impactul, data.gov.ro poate fi proactivă în această speță și poate dezvolta servicii în ambele direcții: creșterea numărului de seturi de date găzduite în paralel cu indexarea seturilor de date furnizate de alte instituții publice prin canale proprii. Mai exact, poate indexa datele furnizate de alții, asociindu-le metadata, și le poate promova prin portalul propriu. Acțiunea (urmată spre exemplu de portalul danez de date deschise) prezintă avantajul de a nu depune efort pentru arhivare și stocare efectivă, dar implică alocarea de resurse pentru indexarea constantă a noutăților. O strategie de minimizare a costurilor și de distribuire a acestora către alte entități o constituie parteneriatul cu acestea. Data.gov.ro își poate asuma responsabilitatea publicității prin includerea în motorul propriu de căutare, iar partenerii își pot asuma sarcina de a furniza actualizări periodice, în formatul standardizat agreat și de a asigura preluarea automată a acestor informații în portalul data.gov.ro.

Această abordare poate implica, de la caz la caz, implementarea unor mecanisme noi de colaborare cu alte instituții publice, dar și modificări ale unor acte administrative sau chiar normative. Este recomandat ca echipa data.gov.ro să inițieze discuții bilaterale și/sau multilaterale cu celelalte instituții publice furnizoare de date, și în primul rând cele indicate de (re)utilizatori ca prioritate.

Apropriata adoptare a Legii datelor deschise, care va impune instituțiilor publice noi obligații în domeniul publicării datelor, este un bun prilej de a iniția astfel de discuții.

DIFUZAREA DE DATE PRODUSE ÎN MEDIUL ACADEMIC

Există o legitimitate medie în cadrul universitarilor de a difuza gratuit, în formă structurată, produsele activității lor de cercetare ce sunt stocate în seturi de date. Există însă și reticența de a plasa sarcina în sfera de activitate a propriului departament. SGG și data.gov.ro pot dezvolta un program complex care să pună în valoare acest tip de legitimitate existent. Pentru aceasta este nevoie de o atitudine proactivă, pentru că simpla oferire a mediului prin care datele pot fi circulate se dovedește a fi insuficientă. Prin urmare, un program în acest sens poate include:

- Leaflet-uri / newsletter-uri de promovare, pe de o parte către producători de date din mediul academic, pe de alta către mediul academic ca utilizator al acestor date.
- Un program pilot prin care datele din 5-6 departamente diferite, din domenii distincte, să fie aduse în spațiul public prin efort conjugat (universități-SGG, sau terți care acționează în numele SGG)
- Training-uri pentru mediul academic care să promoveze termenii și modurile de a pregăti datele pentru deschidere.

Ultimul punct este esențial în a asigura calitatea datelor. Cel de-al doilea poate avea ca impact efecte de contagiune în interiorul universităților unde se pilotează, respectiv a domeniilor pilot. Primul acționează ca promotor al celorlalte două și aduce activitatea programului pilor în atenția celor ce pot replica programul prin efort propriu.

PROGRAM DE POPULARIZARE

Reprezentarea puternic negativă asupra deschiderii datelor în comparație cu restul UE creează o nevoie certă de promovare și popularizare.

Observațiile simple ale experților consultați au relevat absența data.gov.ro din motorul de promovare a guvernului: pe site-ul guvernului și pe cele ale ministerelor și agențiilor centrale, lipsesc linkurile către data.gov.ro.

Aducerea în prim plan a poveștilor de succes, spuse chiar de cei care le-au înfăptuit, poate fi parte a strategiei de comunicare. De notat că aceasta are tipuri diferite de public (spre exemplu potențiali voluntari sau antreprenori, dar și administrația publică locală) și are nevoie de mesaje distincte pentru categoriile de public. Spre exemplu, pentru administrația publică, povestea de succes relevantă nu este cea a unui ONG care reutilizează datele și le aduce apoi din nou în spațiul public, ci cea a unui funcționar public care rezolvă sarcini curente folosind portalul guvernamental de date.

O recomandare importantă rezultată din aproape toate instrumentele utilizate este că tema datelor deschise are nevoie de asumare politică, preferabil la nivel înalt.

PROGRAM DE MONITORIZARE PERMANENTĂ

Mai ales la nivelul experților, dar și la cel al utilizatorilor (studiile de caz), se conturează ideea necesității unui instrument de monitorizare a data.gov.ro. Acesta poate fi asigurat parțial prin chestionarul de satisfacție, prin publicarea constantă a altmetrics. Rapoarte de impact precum acesta pot fi generat o dată la 2-3 ani, prin repetarea instrumentelor folosite în textul de față. Unele instrumente pot fi repetate la costuri mici sau *in house*, cu o frecvență mai ridicată (o dată pe an).

DATA.GOV.RO CA FURNIZOR DE PREGĂTIRE ȘI CUNOAȘTERE

Datele de cercetare arată că în sectorul public se semnalează un deficit de cunoaștere cu privire la utilizarea datelor în general, și a datelor deschise în particular. Unele temeri și rețineri în a utiliza sau furniza date deschise pot fi explicate prin lipsa abilităților necesare acestui domeniu.

Data.gov.ro furnizează deja o serie de ghiduri și instrucțiuni, care au fost uneori menționate, și sunt utile. Realitatea arată însă că aceste materiale de auto-pregătire nu sunt suficiente pentru a atinge obiectivul de formare a persoanelor responsabile cu gestionarea datelor deschise în instituțiile publice.

Adoptarea Legii datelor deschise va pune o presiune suplimentară pe toate instituțiile publice. Dacă până acum, publicarea seturilor de date a fost mai degrabă o recomandare, adoptată prin acte de tip Memorandum sau prin formulări normative generale, în viitorul apropiat vor exista obligații legale explicite și comprehensive.

În acest context, vor crește așteptările de la portalul național de date deschise să ofere semnificativ mai mult decât o platformă tehnică și un minim ghidaj. Pro-activ, echipa data.gov.ro poate să dezvolte un program de

formare și o bibliotecă de cunoștințe, cu care să vină în întâmpinarea acestor așteptări, și să devină un veritabil centru de expertiză și formare în domeniul datelor (deschise), pentru administrația publică și nu numai.

Programe de training în utilizarea datelor, precum și în publicarea acestora ar constitui un element de accelerare al întregului proces, dar și de eficientizare a modului de folosire a datelor.

INSTRUMENTE DE MONITORIZARE

În bună măsură, acest volum a fixat un benchmark pentru monitorizarea impactului deschiderii datelor. Metodele de analiză propuse au fost construite în așa fel încât sunt replicabile la costuri mici. Datele rezultate din sondajele realizate sunt deja disponibile pe data.gov.ro. În plus, recomandăm monitorizarea permanentă prin chestionare de feed-back și analiza accesărilor site-ului, după cum explicăm în continuare.

CHESTIONARE DE FEEDBACK ASOCIAT DATA.GOV.RO

Ideea centrală din spatele chestionarului de feedback este de a avea un răspuns rapid la utilizării imediate a data.gov.ro. Pentru aceasta, metoda cea mai facilă este de a genera un chestionar dedicat utilizatorilor care accesează data.gov.ro. Este vorba de două tipuri distincte de utilizatori:

- cei care doar accesează site-ul și
- cei care descarcă date din data.gov.ro.

Primii sunt importanți ca potențiali utilizatori ce nu au găsit ceea ce le era necesar însă pe data.gov.ro. Cauza poate fi absența datelor din domeniul de interes, formatul în care este prezentat, felul în care este structurat site, faptul că de fapt nu căutau date ci doar vizitau site-ul, etc. Pentru ei se poate genera un chestionar de feedback, dar este puțin probabil să ofere multe răspunsuri.

Cei din urmă, deja au descărcat date, deci au la ce să se refere și este util să fie investigați mai în amănunt.

Pentru ambele categorii se poate proceda la un chestionar comun. Acesta apare după descărcarea unui set de date sau la finalul navigării în pagină. Recomandarea este ca acest chestionar să fie generat folosind o soluție open-source, precum LimeSurvey. Aceasta permite actualizarea chestionarului in-house, generarea în orice moment a unor statistici descriptive, descărcarea datelor în format tabelar, deschis și în diferite alte formate care sunt mai ușor de analizat rapid cu metode mai sofisticate.

Chestionarul propus este descris în Anexa acestui volum, incluzând precizări pentru persoana care va implementa chestionarul în site. În cazul în care este utilizată o soluție precum, cea recomandată (LimeSurvey), multe dintre indicațiile respective sunt deja parte a acestui soft.

Chestionarul conduce la descărcarea unei „baze de date”, în sensul acordat termenului în științele sociale, în epidemiologie, studii medicale, științe comportamentale, economie șamd. Cu alte cuvinte este vorba de o

tabelă, în care fiecare coloană reprezintă răspunsul la o variabilă din chestionar și este etichetată cu un nume unic. Fiecare rând constituie un respondent.

În cele ce urmează, etichetele cu care începe fiecare întrebare reprezintă numele variabilei (capul de coloană în baza de date). Toate codurile din baza de date sunt numerice, exceptând cele câteva câmpuri text specificate ca atare. Codurile numerice asociate fiecărei variante de răspuns sunt precizate înaintea acesteia.

În paranteze drepte apar instrucțiuni pentru designul chestionarului. Ele nu trebuie să apară în ecranul pe care îl vede respondentul. De asemenea, respondentul nu va vedea numele variabilelor sau codurile asociate variantelor de răspuns, exceptând locurile unde am precizat explicit că aceste numere sunt vizibile.

BACK-LOG ȘI ALTMETRICS

Analiza statisticilor legate de utilizare poate fi realizată permanent, în felul în care am construit-o în capitolul dedicat (*Back-log analysis*, pagina 22)

REFERINȚE

- Bice, S. (2020). The future of impact assessment: problems, solutions and recommendations. *Impact Assessment and Project Appraisal*, 38(2), 104-108.
- Blank, M. (2019). *Open Data Maturity Report 2019*, European Data Portal, https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf, accesat 7.07.2020.
- Bond, A., & Dusík, J. (2020). Impact assessment for the twenty-first century – rising to the challenge. *Impact Assessment and Project Appraisal*, 38(2): 94-99.
- Bond, A., & Pope, J. (2012). the state of the art of impact assessment în 2012. *Impact Assessment and Project Appraisal*, 30(1), 1-4.
- Bulle, C., Margni, M., Patouillard, L., Boulay, A. M., Bourgault, G., De Bruille, V., ... & Kashef-Haghighi, S. (2019). IMPACT World+: a globally regionalized life cycle impact assessment method. *The International Journal of Life Cycle Assessment*, 24(9), 1653-1674.
- Carpenter, J., Watts, P. (2013) *Assessing the Value of Ordnance Survey OpenData (OS OpenData™) to the Economy of Great Britain*, UK Government.
- Carrara, W, Radu, C., Vollers, H. (2017). Open Data Maturity în Europe 2017. European Data Portal, https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n3_2017.pdf, accesat 7.07.2020.
- Castro, D., & Korte, T. (2015). Open Data în the G8: A Review of Progress on the Open Data Charter. *Center for data innovations. US*, available at <http://www2.datainnovation.org/2015-open-datag8.pdf> (accessed May 01, 2016).
- Ciolan, L., coord., (2009). Manual pentru evaluarea ex-ante a impactului politicilor educaționale. București: Secretariatul General al Guvernului - Direcția de Politici Publice.
- Corrales-Garay, D., Mora-Valentín, E. M., & Ortiz-de-Urbina-Criado, M. (2019). Open data for open innovation: An analysis of literature characteristics. *Future Internet*, 11(3), 77.
- Davies, T., Perini, F. & Alonso, J.M. (2013). *Researching the emerging impacts of open data: ODDC conceptual framework*. Washington DC: World Wide Web Foundation.
- Deighton-Smith, R., & Jacobs, S. H. (1997). *Regulatory impact analysis: best practices în OECD countries*. OECD.
- Deloitte, (2018). *The Socio Economic Impact of Open ELS*, Warsaw.
- Drăgan, G., & Nicolescu, A. (2009). *Studiile de impact-instrumente pentru o mai bună reglementare a politicilor publice în Uniunea Europeană* (No. 24). EIR Working Papers Series.
- Dreyer, L., Hauschild, M., & Schierbeck, J. (2006). A framework for social life cycle impact assessment (10 pp). *The International Journal of Life Cycle Assessment*, 11(2), 88-97.
- Edmonton City. (2017). *Open data Strategy: Open City*, Edmonton, Canada.
- European Commission, (2020). *The Digital Economy and Society Index (DESI). Thematic chapters*, Bruxelles: European Commission.

- Frank, M., Simperl, E., Walker, J. (2020). *Sustainability of (Open) Data Portal Infrastructures Measuring Use and Impact of Portals*, Luxembourg: Publications Office of the European Union
- G8 (2013). *Open Data Charter and Technical Annex*.
- GRU (Government Reform Unit. Department of Public Expenditure and Reform). 2017. *Open Data Strategy 2017 – 2022*, Dublin. https://data.gov.ie/uploads/page_images/2018-03-07-114306.063816Final-Strategy-online-version1.pdf
- Herala, A. (2018). Benefits from Open Data: Barriers to Supply and Demand of Open Data in Private Organizations. *Acta Universitatis Lappeenrantaensis* 809
- Huijbregts, M. A., Steinmann, Z. J., Elshout, P. M., Stam, G., Verones, F., Vieira, M., ... & van Zelm, R. (2017). ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *The International Journal of Life Cycle Assessment*, 22(2), 138-147.
- Huyer, E., van Knippenberg, L. (2020). *The Impact of Open Data in Economy*, <https://www.europeandataportal.eu/sites/default/files/the-economic-impact-of-open-data.pdf>, accesat pe 7.07.2020.
- Jäppinen, S., Toivonen, T., & Salonen, M. (2013). Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: An open data approach. *Applied Geography*, 43, 13-24.
- Jolliet, O., Margni, M., Charles, R., Humbert, S., Payet, J., Rebitzer, G., & Rosenbaum, R. (2003). IMPACT 2002+: a new life cycle impact assessment methodology. *The international journal of life cycle assessment*, 8(6), 324.
- Kirkpatrick, C., Parker, D., & Zhang, Y. F. (2004). Regulatory impact assessment in developing and transition economies: a survey of current practice. *Public money and management*, 24(5), 291-296.
- Maail, A.G. (2017). The relational impact of open data intermediation: Experience from Indonesia and the Philippines, pp. 153-166 in Francois van Schalkwyk, Stefaan Verhulst, Gustavo Magalhaes, Juan Pane & Johanna Walker (eds). *The Social Dynamics of Open Data*, Cape Town: African Minds.
- Martin, S., Turki, S. (2018). *Impacts of Open Data in Luxembourg and the Greater Region - 2018*, Luxembourg Institute of Science and Technology (LIST).
- Ministry of Information Society and Administration (MISA) (2018). *Open Data Strategy (2018-2020)*, Government of the Republic of Macedonia
- Monino, J. L., & Sedkaoui, S. (2016). *Big Data, Open Data and Data Development* (Vol. 3). John Wiley & Sons.
- OECD (2019), *Rapport sur les données ouvertes publiques: Encourager la maturité des politiques de données ouvertes pour un impact durable*, OECD Publishing, Paris, <https://doi.org/10.1787/12ea5027-fr>
- Open Data Charter. (2015). International open data charter. <https://opendatacharter.net/> , accesat 7.07.2020.
- PwC (2017) *The impact of the open geographical data – follow up study Agency for Data Supply and Efficiency*, <https://sdfe.dk/media/2917052/20170317-the-impact-of-the-open-geographical-data-management-summary-version-13-pwc-qrvkvdr.pdf>

- Radaelli, C. M. (2004). The diffusion of regulatory impact analysis—Best practice or lesson-drawing?. *European Journal of Political Research*, 43(5), 723-747.
- Radaelli, C. M. (2009). Measuring policy learning: regulatory impact assessment in Europe. *Journal of European public policy*, 16(8), 1145-1164.
- Renda, A. (2006). *Impact Assessment in the EU: The State of the Art and the Art of the State*. Bruxelles: Centre for European Policy Studies.
- Rinciog, O., & Posea, V. (2018). State of Romanian Open Data in 2017. *University Politehnica of Bucharest Scientific Bulletin Series C-Electrical Engineering and Computer Science*, 80(4), 3-12.
- Stoneman, J. (2015). Does Open Data Need Journalism?. *Reuters Institute for the Study of Journalism, Working paper*, <https://ora.ox.ac.uk/catalog/uuid:c22432ea-3ddc-40ad-a72b-ee9566d22b97>
- Therivel, R. (2020). Impact assessment: from whale to shark. *Impact Assessment and Project Appraisal*, 38(2), 118-121.
- Thorup, H., Frandsen, K., Sass Larsen, H., Christensen, V., Aaen, F., Bager, B. (2019). *Beretning om åbne data*, København: Folketinget.
- van Knippenberg, L. (2020a) *Analytical Report 16. Open Data Best Practices in Europe: Learning from Cyprus, France, and Ireland*. European Data Portal.
- van Knippenberg, L. (2020b) *Open Data Maturity Report*. European Data Portal.
- Van Schalkwyk, F., Canares, M., Chattapdhyay, S., & Andrason, A. (2016). Open Data Intermediaries in Developing Countries. *The Journal of Community Informatics*, 12(2).
- Vanclay, F. (2020). Reflections on Social Impact Assessment in the 21st century. *Impact Assessment and Project Appraisal* 38(2): 126-131.
- Vasiu, R. (2015, March). Development of Smart City Applications Based on Open Data. In *Contribution to the 13th NETTIES Conference (Network Entities) on "Open Data and Big Data—The Impact on Digital Society and Smart Cities"*. Humboldt Cosmos Multiversity, Tenerife (Vol. 2).
- Verhulst, S. & Young, A. (2016). *Open Data Impact: When demand and supply meet*. New York: The GovLab and Omidyar Network.
- Voicu, B. (2018). *Creșterea calității și a numărului de seturi de date deschise publicate de instituțiile publice. Rezultatele anchetelor pe un eșantion de organizații publice*. Ernst & Young și SGG.
- Weidema, B. P. (2006). The integration of economic and social aspects in life cycle impact assessment. *The International Journal of Life Cycle Assessment*, 11(1), 89-96.
- Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17-29.

CHESTIONAR DE SATISFACȚIE A UTILIZĂRII DATA.GOV.RO

Bună ziua,

Vă mulțumim pentru vizita dvs. pe site-ul nostru. În vederea îmbunătățirii acestuia, dorim să aflăm câteva opinii despre site, printr-un chestionar care durează 5-7 minute. Menționăm că răspunsurile dvs. sunt complet anonime și vor fi utilizate doar pentru statistică.

FILTRU. Ați descărcat vreun set de date azi de pe data.gov.ro

1. da
2. nu

Care dintre următoarele se aplică vizitei dvs. de astăzi:

(Vă rugăm să bifați toate variantele care vi se potrivesc.)

[întrebare cu răspunsuri multiple, adică respondentul poate bifa mai multe variante. În baza de date va fi câte o coloană pentru fiecare tip de răspuns. Pentru fiecare variabilă (i.e. coloană) dacă se bifează varianta respectivă apare în baza de date codul 1, iar dacă nu se bifează apare 0. Pentru VIZ5, dacă este bifat apare un câmp text în care respondentul poate introduce motivul respectiv]

VIZ1. Descărcarea de date

VIZ2. Curiozitatea

VIZ3. Am vrut să văd dacă sunt actualizări ale unor date descărcate

VIZ4. Căutam meta-data pentru date deja descărcate

VIZ5. Alt motiv

[Dacă FILTRU=2 și nu a bifat VIZ1 sau VIZ3, chestionarul se oprește]

[Dacă FILTRU=2 și VIZ1=1 sau VIZ3=1, atunci apar întrebările MN1-MN6, apoi se sare direct la SECTOR1]

Care este motivul pentru care nu ați descărcat date:

(Vă rugăm să bifați toate variantele care vi se potrivesc.)

[întrebare cu răspunsuri multiple, adică respondentul poate bifa mai multe variante. În baza de date va fi câte o coloană pentru fiecare tip de răspuns. Pentru fiecare variabilă (i.e. coloană) dacă se bifează varianta respectivă apare în baza de date codul 1, iar dacă nu se bifează apare 0. Pentru MN6, dacă este bifat apare un câmp text în care respondentul poate introduce motivul respectiv]

MN1. Data.gov.ro nu include datele pe care le caut

MN2. Actualizările căutate nu sunt disponibile

MN3. M-am răzgândit

MN4. Meta-data sunt neclare sau insuficiente

MN5. Site-ul este dificil de navigat

MN6. Alt motiv

[dacă FILTRU=1, MN1-MN6 nu apar, dar se continuă cu restul chestionarului]

În ce măsură sunteți satisfăcut fiecare din următoarele aspecte. Vă rugăm să notați fiecare aspect pe o scală de la 0=total nesatisfăcut la 10=total satisfăcut, putând alege orice valoare intermediară.

[de preferință *slider*, cu capetele 0 și 10, etichetate corespunzător, plus variantele 98=nu știu și 99=nu răspund; pentru „nu știu” și „nu răspund” nu vor apare pe ecran codurile 98 și 99; ele apar doar în baza de date]

SAT1. Disponibilitatea datelor

SAT2. Formatul datelor

SAT3. Ușurința accesării

SAT4. Modul de organizare a site-ului

SAT5. Motorul de căutare

SAT6. Calitatea metadata

SAT7. Posibilitățile de a lăsa feed-back pentru un set de date

În ce sector activați:

(Vă rugăm să bifați toate variantele care vi se potrivesc.)

SECTOR1. Public (de stat)

SECTOR2. Privat

SECTOR3. ONG

SECTOR4. Sunt free-lancer (liber profesionist)

SECTOR5. Sunt elev sau student

SECTOR6. Niciunul din cele de mai sus.

[atenție la design: SECTOR6 nu poate fi bifat simultan cu alte variante. Dacă este bifat, atunci celelalte se debifează automat]

Activați în

(Vă rugăm să bifați toate variantele care vi se potrivesc.)

SS1. Administrație publică

SS2. Business

SS3. Educație

SS4. Cercetare

SS5. Mass-media

SS6. ONG

SS7. Consultanță

SS8. Altele

SS9. Niciuna din cele de mai sus

[întrebare cu răspunsuri multiple, adică respondentul poate bifa mai multe variante. În baza de date va fi câte o coloană pentru fiecare tip de răspuns. Pentru fiecare variabilă (i.e. coloană) dacă se bifează varianta respectivă apare în baza de date codul 1, iar dacă nu se bifează apare 0. Pentru SS9, dacă este bifat apare un câmp text în care respondentul poate introduce motivul respectiv]

În ce scop folosiți datele

SCOP1. Personal

SCOP2. Business

SCOP3. Academic, la școală

SCOP4. Non-profit

[întrebare cu răspunsuri multiple, adică respondentul poate bifa mai multe variante. În baza de date va fi câte o coloană pentru fiecare tip de răspuns. Pentru fiecare variabilă (i.e. coloană) dacă se bifează varianta respectivă apare în baza de date codul 1, iar dacă nu se bifează apare 0.]

EXP. Cum ați aprecia experiența dvs. în folosirea de date? Vă rugăm să folosiți o scală de la 0=foarte redusă la 10=foarte mare, putând alege orice valoare intermediară.

[de preferință *slider*, cu capetele 0 și 10, etichetate corespunzător, plus variantele 98=nu știu și 99=nu răspund; pentru „nu știu” și „nu răspund” nu vor apare pe ecran codurile 98 și 99; ele apar doar în baza de date]

PANEL. Dorim să revenim către dvs. peste câteva săptămâni cu un scurt chestionar despre calitatea datelor descărcate. Dacă sunteți de acord, vă rugăm să ne lăsați aici o adresă de email (va fi eliminată ulterior din baza de date pentru a nu permite asocierea răspunsurilor cu persoana dvs.)

1. da, doresc să răspund și peste câteva săptămâni
2. nu, nu doresc

[dacă răspunde cu 1, atunci apare un câmp text unde va introduce emailul, și vor exista reguli de validare a emailului]

De preluat automat din site: ce seturi de date a downloadat?

Vă mulțumim pentru răspunsurile dvs.!

CHESTIONAR DE FEEDBACK DUPĂ 3 SĂPTĂMÂNI

(pentru cei ce au lăsat emailul)

Bună ziua,

În urmă cu trei săptămâni ați descărcat de pe data.gov.ro unul sau mai multe seturi de date. Ați spus atunci că sunteți disponibil(ă) pentru a ne împărtăși din experiența dvs. de utilizare. Vă rugăm să faceți acest lucru completând următorul chestionar de 3 minute:

FOL. Ați folosit deja datele?

1. Da, am folosit sau am început să le folosesc
 2. Nu, dar probabil o voi face
 3. Nu, și nu o voi face
 4. Nu știu încă
99. Nu răspund

În ce măsură sunteți satisfăcut fiecare din următoarele aspecte. Vă rugăm să notați fiecare aspect pe o scală de la 0=total nesatisfăcut la 10=total satisfăcut, putând alege orice valoare intermediară.

SATx1. Calitatea datelor

SATx2. Prezența metadata

SATx3. Calitatea metadata

[de preferință *slider*, cu capetele 0 și 10, etichetate corespunzător, plus variantele 98=nu știu și 99=nu răspund; pentru „nu știu” și „nu răspund” nu vor apare pe ecran codurile 98 și 99; ele apar doar în baza de date]

În ce măsură vi se aplică fiecare din următoarele

[un răspuns pe fiecare rând]

		În foarte mică măsură	În mică măsură	În mare măsură	În foarte mare măsură	Nu pot aprecia	Nu răspund
FF1	folosiți datele în forma preluată de pe data.gov.ro	1	2	3	4	98	99
FF2	e nevoie de pregătire a datelor pentru folosire	1	2	3	4	98	99
FF3	folosiți datele complementar cu alte seturi de date	1	2	3	4	98	99
FF4	Setul de date are erori	1	2	3	4	98	99
FF5	Setul de date este de încredere	1	2	3	4	98	99
FF6	Ar fi fost utile mai multe informații	1	2	3	4	98	99
FF7	Reutilizez setul de date pentru a genera date noi	1	2	3	4	98	99

OPEN. În datele pe care le-ați descărcat de pe data.gov.ro, ce elemente au lipsit sau puteau fi îmbunătățite?

(Vă rugăm să notați cât mai mult. Feedback-ul dvs. ne este de mare ajutor!)

..... [câmp text mare, poate scrie ce vrea aici]

Vă mulțumim pentru răspunsurile dvs.!

CE SETURI NOI DE DATE SUNT SOLICITATE PE DATA.GOV.RO

Lista de mai jos include răspunsurile primite de la cei ce au completat chestionarele sondajelor exploatate în acest volum. Le reproducem ca atare, păstrând grafia originală și prezentându-le într-o ordine aleatorie.

- date colectate de la senzori ce privesc atmosfera - date statistice privind
- date referitoare la ceea ce se produce efectiv în România, cât se importa (c
- Informații din toate domeniile acoperite de aparatul guvernamental
- orice evoluție temporală de date ce pot fi interpretate prin intermediul unor modele de dinamică a sistemului respectiv - format ASCII sau txt, orice format importabil de date
- procedurile de lucru, necesare pentru controlul intern managerial. Pentru o
- PROIECTELE EUROPENE ȘI REZULTATELE ACESTORA - FINANȚĂRILE ONG-urilor ȘI FUNDAȚIILOR CU STATUT DE UTILITATE PUBLICĂ
- Seturi de date spațiale (corelare cu portalul INSPIRE) - date în format mai
- seturi de date statistice produse la nivel de UAT,
1. cataloagele muzeelor. 2. cataloagele bibliotecilor. 3. calendarul eveniment
academice
acces publicatii fara taxa
Acolo unde este cazul, datele tabelare să fie interconectate cu un geoportal (de tipul geoportal ANCP), sau cum am văzut la raportările ANM pe data.gov.ro și date relaționate geografic
anchete sociologice, etnologice etc.
Ar trebui să poată fi accesate toate datele, mai puțin cele secretizate. Trebuie să existe un loc credibil și validat de statul român, de unde poți extrage ușor informațiile necesare.
Articole
Aș dori să găsesc date descărcabile despre populație (la nivel infraurban), economie (date fiabile, cantitative, pe ramuri ale industriei - producții fizice, valori etc. - producții suprafețe și producții agricole, tipuri de servicii etc.), agregate de la nivel local, până la nivel național. Desigur, firesc ar fi să existe pe site-ul Institutului Național de Statistică, dar din păcate nu există. Statistica din alte state europene este mult mai fiabilă, cu date defalcate până la nivel de microcartier (pentru populație, de exemplu), foarte utile pentru cunoașterea și managementul performant al teritoriului, al localităților, pentru realizarea de strategii de dezvoltare și de politici publice care să răspundă nevoilor populației. Problemele de planificare a dezvoltării teritoriale își au deseori sursa în lipsa datelor, în imposibilitatea de a cunoaște, evalua și prognoza dezvoltarea, evoluția societății socială, pentru că nu există date, decât la nivel de regiune (unele), județ (multe) sau localitate. Sunt și date care nu există decât la nivel național (economie, calitatea vieții etc.), ceea ce este aberant, pentru că nu se pot face nici un fel de analize pertinente, utile societății românești. Chiar și datele în acces liber de la ultimele recensăminte sunt mult sub nivelul de prelucrare și sub forma de prezentare a datelor recensământului din 1930!!!, chiar dacă acestea din urmă sunt doar în format tipărit. Este aberant ca, într-o societate deschisă, cu mijloacele digitale actuale, să nu putem dispune de un grad mai mare de detaliere și fiabilitate a datelor despre populație, gospodărie, locuințe etc.!!! Până și un indicator grosier (de bază), cum este numărul populației rezidente, este disponibil doar la nivel de județ!!! Probabil că nu toate aceste tipuri de date trebuie să fie postate pe site-ul gov.ro, dar ele nu sunt accesibile nicăieri. Ar trebui să fie accesibile la Institutul Național de Statistică și la direcțiile regionale/județene de statistică, dar nu sunt!

Atat date cu caracter macroeconomic, cat și cu caracter microeconomic.
Barometre de opinie
Baze de date rezultate din proiectele de cercetare
Beneficiarii finanțărilor din fonduri publice
buget, statistici de tot felul
Bugete alocate, execuții bugetare corespunzătoare planurilor naționale/ str
Bugete strategii, Programe, Monitorizare, implementare. Acum am intrat și am v
bugetele și cheltuielile din toate institutiile publice
Câștigurile actorilor internaționali ce activează în România și proporția folosirii lor pentru (1) investiții în România, (2) scoatere din România; Situația detaliată a furnizorilor de „șpăgi” - cel puțin pentru dosarele penale în care români au fost condamnați pt primire de asemenea sume de bani sau alte avantaje; Situația detaliată a pădurilor României și a tăierilor din acestea, cu detalieri privindu-le pe cele virgine. Situația detaliată a evoluțiilor numărului de personal administrativ concediat și, respectiv, angajat, pt fiecare legislatură (la nivele centrale și locale) și salariile aferente; Câștigurile angajaților la stat cu salarii peste 5000 lei și performanțele acestora în respectivele joburi (ce au făcut ei, detaliat și ce rezultate au avut cele făcute de ei, utile și inutile pt anjatori). Situația centralizată a acțiunilor administrative (ale diverselor instituții, începând cu cele guvernamentale) la diversele probleme apărute în țară: „viteza de reacție”, resursele mobilizate (fonduri, număr de personal, aparaturi etc.) și efectele rezultate (pagube materiale, pierderi de altă natură, inclusiv de vieți, pe zile).
cele de interese public
cele definite intr-un cadrul institutional
cele oferite și de organizatia mea...ma refer strict la directiile din ministe
centralizări, situații statistice etc
Cercetare IT
Cercetari stiintifice!
Coduri caen și clasificare activități
consumul de medicamente din România , pe clase farmaceutice (coduri ATC) consumul de medicamente eliberate cu rețetă/fără rețetă din farmaciile deschise/cu circuit închis
Dat fiind că nu am folosit dat.gov.ro până acum, mi-e greu să știu exact ce fel de date mi-aș dori să fie publicate aici.
data privind densitatea traficului urban, date privind dinamica populatiei (nasteri/ decese), date privind regimul pluviometric în diverse locatii specifice (in special zone agricole)
data privind salarizarea din sectorul public
Date administrative dezagregate la nivel de localitate, dupa cod SIRUTA care sa permita analiza individuala nu la nivel agregat.
date aferente diferitelor echipamente pentru cercetare
Date ale unor studii statistice din domeniul medical din Romania, realizate în profesioniști din mediul academic.

Date brute pentru indicii calitativi ai produselor alimentare
Date care ar putea fi utilizate în cercetare, atât în format tabelar, cât și imagini multi-dimensionale, de genul imaginilor sau secvențelor de imagini (de exemplu în sanatare, pentru metode de diagnostic automat).
Date care din categoria indicatorilor social-economici și de dezvoltare durabilă la nivel de localitate, în format Xcel. Date spațiale utilizabile de softuri de specialitate pentru analize privind dinamica teritorială: evoluția categoriilor de utilizare a terenului, a suprafețelor împădurite, a suprafețelor construite etc....a terenurilor cu riscuri de diferite tipuri.
Date care pot fi folosite în activitatea de cercetare pe domeniile în care activez sau în domeniile pe care doresc să mi dezvolt abilități.
Date care rezultă din Politicile UE (Directive, Regulamente), Date despre Programe de Finanțare/Finanțatori; Date despre constituirea de Consorții de Cercetare-Dezvoltare în domenii ca: Siguranță Alimentară, Securitate Alimentară, Toxicitatea sau Inocuitatea Produselor Alimentare (care să vizeze domenii de genul Listele de Aditivi alimentari permisi și dozele aferente pentru cei permisi în diverse state UE, non-UE, internațional); Date privind Inspecția, Certificarea, Acreditarea (Voluntară/Reglementată) a unor Produse/Servicii; Date privind Politicile Publice ale României pe diverse Orizonturi...
Date care să poată fi utilizate în activitatea de cercetare, în activitatea didactică și în activitatea de publicare de lucrări științifice.
Date climatice de la stațiile meteorologice platite de guvern, dar revândut pe bani grei de ANM, chiar dacă sunt folosite exclusiv în scop didactic și științific
date climatice, date biodiversitate, date economice pe regiuni, județe (cel puțin)
date cu formulări clare, logică în dezvoltarea și accesarea paginilor și în clarificarea mesajului, interconectabilitate cu alte site-uri de pe data.gov, sursa din care se susțin datele (de exemplu referiri clare și nu însirui de surse alambicate și fără legătura directă cu datele comunicate. "Nu vreau calitate"
date cu privire la fiecare organizație a administrației publice (statistici re
Date de interes public.
date de interes strategic (referitoare la diverse diagnosticări efectuate sectorial, tendințe structurale pe termen mediu și lung, schimbări radicale preconizate în economie etc)
date de natură științifică
date de preferință declarată ale utilizatorilor de transport public de călători (în transport urban și periurban/local), în serii de timp
Date demografice pe localități Date tip household survey pe localități Date de consum pe localități Date de producție internă agricolă, mai ales producători mici/locali Date de accesare fonduri la producătorii locali
date demografice, date legate de epidemiologia unor patologii la om și animale, date rezultate din diverse studii populationale
Date despre apele uzate
Date despre calitatea apelor din România din punct de vedere fizico-chimic și biologic Liste de specii de organisme care trăiesc în râuri și lacuri Baze de date cu parametri fizico-chimici ai apei râurile și lacurilor, de exemplu: pH, oxigen dizolvat, nutrienți etc

date despre cati imigranti sunt inregistrati în romania
Date despre cultele și asociatiile religioase
Date despre e-Health în Romania
Date despre economie
Date despre educație atat ca sistem de predare și cercetare, ca retea de conectare între instituțiile de educație.
date despre sanatate, economice, financiare, cercetare, invatamant, sociale
Date despre sănătatea populatiei
date despre societatea romaneasca, categorii de varsta, sociale, venituri, mobilitate
Date despre strategia de învățământ universitar din România.
Date despre urbanism.
Date despre evoluția tuturor ramurilor economiei și legislația în forma finală
date detaliate privind prețurile
date din domeniul cercetării medicale - anonimizate, date de sănătate publică - anonimizate
Date din sectorul agriculturii: - date legate de producții la ha, maxime, medii, la nivel național, regional, județean, la nivel de comună - date legate de suprafețe cultivate cu diferite culturi, - suprafețe exploatate în proprietate / arendate - evoluție prețuri produse agricole - importuri / exporturi - nivelul procesării - date privind irigațiile din sectorul agricol - date privind evoluția economică a unităților cu profil agricol
date economice, financiare, date privind resursele umane
date economice, sociologice
Date epidemiologice și medicale la populația pediatrică.
date experimentare care să cuprindă condiții de desfășurare a experimentelor
date financiare, date din administrația publică (trafic, poluare, rețele informatice, rețele utilități (electrice, apă, gaz), rețele de transport public)
Date fizico-geografice, socio-economice, la nivel național în format csv, txt,
Date geo-spatiale
Date geospațiale (formate .shp, .csv, .xls, .tif) despre soluri (profile de sol georeferențiate, studii pedologice, hărți pedologice de detaliu, etc). Date geospațiale despre utilizarea terenurilor în România. Date geospațiale (limite de județ, UAT, etc) cu atribute actualizate și diverse.
date istorice meteorologice de la ANM, date satelitare și GIS din proiectele m
Date istorice relevante
Date în ceea ce privește populația carcerală din România Statistici cu pr
Date la nivel de UAT care nu sunt acoperite de Tempo.
Date la nivel mezoeconomic, macroeconomic, pe regiuni cu privire la economia României.

Date la zi despre situația demografică a țării. Date la zi despre diverse situații de urgență. Acestea ar fi necesare pentru elaborarea de modele științifice.
Date legate de comportamentul consumatorului în mediu online
Date legate de domeniul educational.
Date legate de învățământul superior: licență, master, doctorat - număr studenți intrare, nr. absolvenți, repartizare pe domenii etc.
Date legate de poluare, statistici legate de sănătate, informații legate de vreme pentru perioade lungi de timp (10-15 ani)
date legate de procentul de conectivitate la internet, rețeaua de electricitate
date medicale
date mediu; date poluare; date consumuri energetice; demografice; sănătate publică; resurse
Date meteorologice
date meteorologice de la ANM, date statistice la nivel local, harta digitală a
date meteorologice de la stațiile meteo pentru ultimii 100 de ani
date necesare întocmirii diferitelor situații centralizatoare
date obținute folosind banii publici.
date pe partea de dezvoltare durabilă
Date preluate de la ANAF
Date preluate de la senzori de tip IoT. Date statistice. Date textuale adnotate de experți.
Date privind aspecte economice, de mediu, indici demografici etc.
date privind dezvoltarea regională la nivel de județ, regiune macroeconomică și regiune istorică, ba chiar și la nivel de localități din județ privind administrația publică ca și ramurile economice pe grupe și subgrupe
Date privind educația și cultura în România, numărul de creșe, școli, universități, gradul de educație al populației, infrastructura culturală (muzee, insituri de spectacole, biblioteci etc), informații privind promovarea, prezervarea și stimularea culturii și artei românești, date și idei privind promovarea și învățarea limbii române și a istoriei poporului român, atât din perspectiva istorică cât și în context actual. Date privind cetățenia europeană și mijloacele particulare ale românilor europeni și minorităților din România, date privind regiunile țării...
Date privind performanța elevilor și școlilor - testul PISA, capacitate, bacalaureat; date privind rezultatele din cercetare ale autorilor din România
date privitoare la libertatea religioasă și regimul juridic al ei, în corelare cu datele la nivel european.
date privitoare la piața muncii la nivel local, județean, regional, național, internațional
Date provenind de la IGSU cu privire la hazardurile produse în Ro - fișiere xls, dxf; Date specifice recensământului; Date cu privire la investițiile specifice energiilor verzi; Date specifice obiectivelor de patrimoniu din Ro: starea acestora, localizare, istoric, importanță
date referitoare la evidența patrimoniului național

Date referitoare la impactul educatiei online supra elevilor, date referitoare la ceea ce functioneaza bine și prost în scoala romaneasca etc
Date referitoare la monumente istorice, protectia și reabilitarea monumentelor, urbanism, turism
DATE REFERITOARE LA SCIM
Date referitoare la tineri: nr angajati, nr. someri, clase de studii, nr. de O
date referitoare la: - educatia în invatamantul superior, - absorbtia pe piata muncii a absolventilor din domeniul IT, - reconversia profesionala catre domeniul IT.
date salariale din administratia publica, dar și din sectorul privat
date salariale din cadrul sectorului public
Date spatiale (referitoare la UAT, orase, comune, cartiere etc.)
date spatiale cu componente statistice
Date specifice domeniului agricol, economic, sanatate, ducatie.
Date specifice firmelor din Romania care pot fi preluate din situatiile financiare depuse.
Date statistice despre diverse categorii sociale din România și diaspora, legate de educație, diversitate culturală
date statistice din domeniul alimentar, agricol, educatie, cercetare, date privind siguranta alimentara
date statistice din industrie, cercetare, sanatate publica, demografie, etc.
Date statistice pe zona de protecție sociala
Date statistice privind demografia și economia
Date statistice referitoare la economie, demografie, cultura în profil general și regional.
Date statistice, actuale privind domeniul agricol .
Date statistice, pe diferite domenii social, economic și chiar stiintific.
date statistice: productia de otel, cupru, energie electrica etc evolutia productiei; export, import, piata de capital
Date statistici
Date stiintifice, de cercetare și statistici institutionale.
Date științifice, pentru cercetare/educație și statistici legate de instituții, date legate de achiziții, investiții, etc
Date toxicologice.
date vectoriale referitoare la UAT-uri, localitati,
Datele referitoare la autorizarea structurilor de cazare/alimentație, agenții de turism etc. http://turism.gov.ro/web/autorizare-turism/
De exemplu, date privind furnizorii de formare profesionala autorizati care po
Debite ale cursurilor de apa Temperaturi Volume în principalele lacuri de acumulare
definitia prezentata este prea generala, fara a fi concretizata în exemple cla

deja am trecut de 10 minute....orice sondaj trebuie sa respecte acest criteriu....
despre mediul de afaceri (pe CAEN, localitati(siruta superior) firme sustinute prin FSE vs firme fara granturi, forta de munca în firme create cu sprijin FSE vs firme fara granturi), despre sanatate (starea de sanatate), despre administratie (cheltuieli, venituri pe categorii), infrastructura
desre locuitorii României, datele din recensământe, date despre producția de cărți, date despre venituri, date despre ziare/periodice, date despre istorie
diferite statistici și informatii economice
diverse
diversi indicatori economici la nivel local (UAT-uri)
diversificarea seturilor de date, a tipurilor/temelor abordate
Domeniul Transporturi: - Administratori de infrastructuri de transport (diferite moduri, diferite niveluri regionale/spațiale): Indicatori financiari pentru activități de investiții și întreținere a componentelor infrastructurii; Volume ale fluxurilor de trafic pe entități (în format dezagregat) ale rețelei - Operatori de transport de călători și mărfuri, la nivel local și național, diferite moduri de transport: Volume ale fluxurilor de transport pe entități (în format dezagregat) ale rețelei la nivel național, regional, local; Indicatori tehnici și de exploatare
Eu am fost interesata de programele nationale de transplant renal, transplant hepatic și transplant de cord. Am gasit ceea ce am cautat la momentul respectiv.
evaluarea culturi organizaționale (folosind instrumentul OCAI) în cadrul instituțiilor publice furnizoare de servicii sociale din România
executia bugetara de la fiecare institutie, obiectivele fixate și masura indep
executiile bugetare anuale la nivelul fiecărei primarii și consiliul judetean din ro
Exemple: Microdate Barometrul de Consum Cultural, Ministerul Culturii Autoturisme proprietate personala pe UAT, cod SIRUTA Rata de penetrare a internetului în gospodarii ale populatiei, cod SIRUTA si, în genere, orice date care stau la baza studiilor sau cercetarilor comandate de institutii publice, finantate din bani publici.
Finante
fondul de cercetare (sumele) institutiilor
format csv masuri privind impactul diferitelor masuri/politici publice
Gradul de satisfacție al universitarilor români în raport cu accesul la diverse date sau publicații științifice internaționale. Gradul de satisfacție al studenților români în raport cu accesul la diverse date sau publicații științifice internaționale din domeniul lor de pregătire.
imagini, sunete, seturi de date pentru reinforcement learning
in domeniul profesional legate de cadrul juridic, cadrul normativ și de standardizare , EXPLICITAREA unor documete preluate si asumate de Romania, Utilizarea specialistilor din domeniile profesionale în realizarea cadrului juridic, normativ și de standardizare, intr-o masura mai mare decât decât a celor din Agentiile statului și oamenii politici
in niciun caz datele personale
Incidenta unor afectiuni medicale în populatie. Incidenta unor fenomene medicale în populatie.

indicatori de performanta pentru institutii/organizatii/ publice
Indicatori economici
Indicatori economici, indicatori demografici, indicatori din doeniul sanatatii
Indicatori relativ la COVID19, la nivel de UAT, pe zile. Indicatori relativ la educație (elevi, profesori, resurse), la nivel de unități școlare și clase. Indicatori relativ la veniturile indivizilor, istoricul educațional și de muncă, la nivel individual, anonimizat (sau măcar UAT). Datele de recensământ la nivel de sector de recensământ sau chiar micro-data. etc.
Indicatori sociali.
Industrie, Robotica, Tehnologie, Societate, Comerț, Divertisment, Turism.
Informatii
informatii de natura financiar-fiscala
Informatii despre numarul de elevi din anii finali ai invatamantului liceal, despre numarul de locuri din uiversitatile de stat și din cele particulare, despre numarul total de cadre didactice din diferitele cicluri ale invatamantului romanesc, despre numarul total de studenti, defalcat pe specialitati și institutii de invatamant superior.
informatii din sport, educatie, business date statistice, numar de participanti, profilul clientilor (pe diferite domenii- de ex al consumatorului de fitness etc) date statistice despre incidenta afectiunilor (de ex: obezitate- infantila) - cu posibilitatea de a manipula datele (ex din mediu rural/urban; cu venituri mici/medii/mari)etc eventual cu posibilitatea de compara și cu date similare la nivel european (!)
Informatii microeconomice (despre firme, consumatori) și macroeconomice (investitii publice, cheltuieli publice, efectele investitiilor și a cheltuielilor publice)
Informații despre educație și rezultatele evaluărilor din școli.
Informații privind mobilitatea populatiei, date cu caracter social la nivel de cartier.
informații privind politica ministerului de externe a României, a raporturilor înterguvernamentale, a raporturilor cu alte organisme și structuri de politică internațională, obligații ale României față de terți cu referire la obiectul obligațiilor...
Institutul national de statistica
invatamant universitar
În general toate acele date relevante pentru a-mi putea forma o opinie corect
La secțiunea Sănătate, care mă interesează pe mine, există un singur set de date. Pentru ca mediul academic să poate face ceea ce trebuie să facă, anume să analizeze date și să ofere feedback, ar fi nevoie de date cu adevărat relevante, pe perioade mai lungi de timp, care să cuprindă indicatori care chiar să vorbească despre performanța sistemului de sănătate.
legate de cercetari în geopolitica
Legislatia nationala gestionata de Monitorul oficial, personal cred ca ar treb
Legislație actualizată, practică judiciară, cărți juridice în format electronic
liste primarii pe judete liste consilii judetene date contact misiuni diplomat
M-ar interesa date legate de problematica minorităților naționale, precum și cele legate de politica bi(multi)lingvismului în România.

Ma gandeam la rezultatele cercetarii romanesti, dar probabil s-a dorit a publica altceva,
Mai multe date despre accesul la educație, la nivel universitar și preuniversitar.
masuratori din centrale electrice regenerabile (PV, vant etc.) masuratori privind consumurile intr-o cladire (cat mai detaliat în ceea ce priveste tipul receptoarelor, caracteristicile cladirii, particularitatile ocupantilor etc) costuri ale echipamentelor și aparatelor electrice de inalta, medie și joasa tensiune (transformatoare, intreruptoare, separatoare, descarcatoare etc.) costuri ale echipamentelor de automatizare ale unei cladiri
Măsurători de calitatea mediului (apa, aer, sol din fiecare localitate din țară, actualizate la 6 luni) Coordonate geografice și de contact pentru toate fabricile din Romania (incluzând cariere, materii prime, halde de depozitare, deșeuri - localizare și cantități) O lista de potențiali parteneri economici (pe domenii) la care studenții ar putea realiza practica tehnologica, internship-uri etc.
Măsurători experimentale, studii clinice
Medicale.
modele de bune practici în domeniul mediului, educatiei, sanataii; sondaje de
N/A
NA
Necesare/utilizate în proiectarea și monitorizarea politicilor publice în domeniile social, de invatamant și cercetare, de sanatate, cultural.
NIMIC
Nimic în afara datelor propuse prin programele guvernamentale.
Nu am idee despre domeniu, nu ma pot pronunta
nu am solicitari specifice
Nu cred ca datele stiintifice sunt de interes pentru opinia publica romaneasca. Datele le public în reviste Internationale de date de tip open source.
Nu cunosc data.gav.ro nu am folosit nici o data nu am auzit pana în acest moment de acest site.
Nu funcționează (nu există date de la) Administrația Fondului Cultural Național!
nu m-am gandit
nu stiu
Nu stiu ce se publica acolo.
Nu stiu, eu lucrez cu corpusuri de limba, cu ceva statistici. Poate ar fi oportun sa existe cataloage de carti și articole stiintifice, clasate pe domenii, articole și carti online, etc.
Nu știu
Nu știu! Date utile
Orice date considerate de interes/utilitate publica.
Orice domeniu de activitate
Orice set de date public poate fi de folos unui cercetator, la un moment dat, cum cercetarea se face multidisciplinar (sau ar trebui - în multe cazuri). Date de natura economica, geografica, demografica,

culturala etc. Exemple: date privind teritoriul: intravilan vs extravilan, date meteo: nivel precipitatii pe zone, date culturale: nr evenimente organizate pe judete, etc, etc
prevederi legale în domeniul de interes fonduri europene disponibile proiecte de accesat
Probleme legate de invatamant și cercetare, de posibilitati fe imbunatire a acestuis
Proceduri de sistem / Proceduri operationale / Instructiuni de lucru specifice
Profesionale, care imi trebuiesc în activitate, functie de specificitate
Publicarea contractorilor care nu si-au respectat clauzele contractuale. Lista
Rapoarte financiare a le societatilor comerciale și ale institutiilor publice
rapoartele de activitate
referitoare la educație - rezultate prelucrate statistic de la examene (de exemplu, de bacalaureat, de admitere la facultate, unde este cazul)
referitoare la: combaterea coruptiei, combaterea criminalitatii, liste cu firme și persoane publice cu probleme de natura fiscala și juridica, legislatie, recensamant
Regulamente Statistici utile Transparenta decizionala Bune practici
Rezultate majore în cercetare.
Rezultate stiintifice ale proiectelor finantate prin fonduri
seturi careau legatura cu managementul riscului, date cu privire la probabilitatile de aparitie a diferite incidente de securitate
Seturi de date ce includ interactiuni de diferite tipuri intre membrii societatii cu scopul de a genera retele ce pot oferi posibilitatea analizei interactiunilor și cum acestea influenta dinamica sistemului social: transmitere de boli (COVID), analiza clasei politice (la nivel de legaturi dintre politicieni și partidele din care fac parte) etc.
Seturi de date legate de activitatea din invatamantul universitar. Seturi de date legate de activitatea de cercetare.
Seturi de date referitoare la cele realizarea de către diferite județe, orașe sau comune a celor 17 obiective de dezvoltare durabilă ai ONU. Seturi de date referitoare la programele de cooperarea transfrontalieră și transnațională.
Seturi de date rezultate din cercetare în domeniul psihologiei, asistentei soc
seturi de date spatiale aferente serviciilor de apa, canalizare, reatea electr
seturi de date specifice
seturi de date stiintifice - nu stiu daca fac obiectul platformei
seturi de indicatori de performanță în universitățile din România informații structurate privind vârsta și genul managerilor din diverse insituții publice informații structurate privind presiunea factorilor de stress asupra performanței în muncă a angajaților din instituții publice
Site-ul ar putea prelua programele și datele publice ale cercetarilor stiintifice care acum se afla pe diferite site-uri în universitati. Astfel le-ar creste vizibilitatea și stabilitatea (adica sa nu le modifice cineva nedorit).
SOFTURI PT.MEDIUL UNIVERSITAR și CERCETARE în INGINERIE ELECTRICA și CALCULATOARE CU LICENTE

software profesional de top, cateva pentru fiecare domeniu de cercetare
sondaje de opinie, statistici privitoare la aspecte sociale, politice, culturale de interes național, regional
standarde tehnice
statistică economică
statistici
Statistici asupra rezultatelor școlare la universități
Statistici Date adnotate, în domenii precum IoT, medical, agricultura
statistici de evidenta a populatiei cat mai granulare inregistrari meteo granulare (la zi, ora) pentru fiecare regiune din tara inregistrari ale dateleor de poluare date din trafic (numar de masini pe anumite tronsoane în anumite intervale) bugetele și contabilitatile publice (guvern, primarii, alte institutii), actuale dar și istoric
statistici despre populație: vârstă, sex, ocupați, loc de muncă, localități, nivel de salarizare informații despre educația preuniversitară: număr de elevi, note, olimpiade, premii, burse informații despre educația universitară: număr de studenți înscriși, note obținute, absorbție pe piața forței de muncă
Statistici legate de sanatate - cum ar fi incidenta bolilor; corelare decese - boli; tratamente medicale - output, ... mediu - poluare - inclusiv corelare cu sanatatea publica; incidenta unor poluanti uzuali/specifici, alimentatie și siguranta alimentara - inclusiv corelare cu sanatatea publica; bp;o specifice, ...
Statistici pe 3, 4, 5 sau 10 ani legate de suprafete, productii, clima, etc.
statistici populationale, statistici starea de sanatate, frecventa bolnavirilor, statistici angajabilitate, piata muncii
statistici urbane, statistici demografice, statistica economica
statistici, informatii actualizate
Statistici, rapoarte, proiecte de legi, diverse acte normative, linii de finantare, programe, etc.
stiintifice, didactice
stiintifice, economico-sociale, medicale
Strategiile guvernamentale și rezultatul consultărilor avute în elaborarea
Toate datele plătite deja de către contribuabilul roman prin plata salariilor angajatilor la stat. De exemplu: informatii cadastrale, datele INS, amenajamente silvice, datele de la ANIF, datele cu caracter public în general.
Toate datele statistice generate în activitatea curentă a instituțiilor pub
Toate domeniile, cat mai multe date, pe cat posibil cu valori din mai multi ani.
Toate seturile de date produse de ministerele din România, în principal ministerul educației.
toti indicatorii de mediu relevanti, implementati gradual, defalcati obligator
totul despre drepturile autor în artele vizuale. orda (de stat) și visarta (de drept privat).... Date despre profesiunea libera de artist vizual în Romania. (observati ca nu prea sunt) și într-o buna zi ,cand va exista și în Romania Legea statutului artistului vizual (promotor sculptorul Dumitru Serban , material elaborat

de 20 de ani,). Atentie e gol legislativ pentru profesia noastra și inca ne confunda cu mesterii populari mai rau decat au facut-o sistemele proletcultiste.
Turism
turism, horeca, marketing și comert
Turism, management
valori ale capacitatilor de productie pentru diverse produse, valori ale capacitatii de prelucrare a materiilor prime, consumuri specifice, costuri



ISBN: 978-606-37-1852-6